

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

EM nucleic acid search, using sw model

Run on: May 14, 2002, 19:29:04, Search time 428.78 seconds
(without alignments)
8400.780 Million cell updates/sec

Title: US-09-445-223-2
Perfect score: 2998
Sequence: 1 gcccattatggtggtatgga.....aaaaaaaaaaaaaaaaaaaaa 2098

Scoring table: M100_NUC
Gapop 60.0, Gapext 60.0

Searched: 1726476 seqs, 953457221 residues

Word size: 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: N_Geneseq_0120021*

- 1: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1989.DAT*
- 2: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1990.DAT*
- 3: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1992.DAT*
- 4: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1993.DAT*
- 5: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1994.DAT*
- 6: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1995.DAT*
- 7: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1996.DAT*
- 8: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1997.DAT*
- 9: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1998.DAT*
- 10: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1989.DAT*
- 11: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1990.DAT*
- 12: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1991.DAT*
- 13: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1992.DAT*
- 14: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1993.DAT*
- 15: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1994.DAT*
- 16: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1995.DAT*
- 17: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1996.DAT*
- 18: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1997.DAT*
- 19: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1998.DAT*
- 20: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA1999.DAT*
- 21: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA2000.DAT*
- 22: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA2001.DAT*
- 23: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA2002.DAT*
- 24: /net/atss06/S1/S1/seqdata/hold/geneseq/geneseq.pseemb1/NA2003.DAT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

Result No.	Score	% Match	Query Length	DB ID	Description
1	2098	100.0	2098	20	AA02558 Human B1 cDNA, full length
2	1769	84.3	2024	21	AA216143 Human cDNA sequence clone
3	1654	78.8	2709	21	AA177779 Human cancer assoc
4	1632	77.8	2033	22	AA094554 Human full length
5	1617	77.1	1931	20	AA209246 Human CARD 3 cDNA
6	1566	74.6	1931	22	AA130001 Human CARD 3 cDNA
7	1518	72.4	2024	21	AA248760 Human RICK coding
8	459	22.3	740	22	AA092259 Human cDNA 5' end
9	468	22.3	740	22	AA093524 Human cDNA clone 1

10	450	16.7	463	21	AA021841 Human breast and o
11	210	10.0	210	19	AA044547 Clone 156108 cDNA
12	171	8.2	575	22	AA094010 Human cDNA 3' end
13	49	2.3	108	16	AA119776 Human gene signal
14	38	1.8	345	22	AA064725 Novel human polyp
15	36	1.6	2082	21	AA079976 Human secreted pro
16	38	1.8	3406	21	AA087701 Human secreted pro
17	38	1.8	3409	21	AA087681 Human secreted pro
18	38	1.8	4062	22	AA044447 Human colon cancer
19	37	1.6	214	22	AA020338 cDNA encoding for
20	37	1.8	216	22	AA062412 Human immune/haem
21	37	1.8	396	22	AA187587 Human polynucleoti
22	37	1.8	427	22	AA061470 cDNA encoding a hum
23	37	1.8	345	22	AA188908 Human polynucleoti
24	37	1.8	443	21	AA059994 Human secreted pro
25	37	1.8	470	22	AA182213 Human polynucleoti
26	37	1.8	351	21	AA050867 Human secreted pro
27	37	1.8	386	18	AA191714 Rat SMIL/CRP2 cDN
28	37	1.8	380	21	AA041708 SMIL/CRP2 related
29	37	1.8	880	22	AA180427 Rat smooth muscle
30	37	1.8	366	20	AA084947 Human secreted pro
31	37	1.8	1125	20	AA030319 cDNA encoding a hum
32	37	1.8	1048	20	AA051370 Human secreted pro
33	37	1.8	1965	20	AA040407 cDNA encoding a hum
34	37	1.8	1602	22	AA056495 Human cDNA for an
35	37	1.8	2020	20	AA070658 cDNA encoding a hu
36	37	1.8	2197	21	AA098120 Human colon cancer
37	37	1.8	3425	21	AA078122 Human cancer assoc
38	37	1.8	3807	20	AA037743 Human PR0326 DNA f
39	37	1.8	4353	20	AA052266 Protein PKC25 cDN
40	37	1.8	1053	21	AA072298 Human PR0326 gene
41	37	1.8	1053	21	AA051115 Human cDNA sequen
42	37	1.8	1053	21	AA051115 Human cDNA clone 1
43	37	1.8	4353	22	AA072424 Human PR0326 cDNA
44	37	1.8	1053	22	AA051115 Human PR0326 polyp
45	37	1.8	1053	22	AA051115 cDNA encoding for

ALIGNMENTS

RESULT 1
AA02558
ID: AA02558 standard cDNA 2098 BP.
XX
AC: AA02558;
XX
XX
BI: 37 MAY 1999 (first entry)
XX
DE: Human B1 cDNA.
XX
ZW: B1 protein, intracellular mediator, modulator, inflammation, cell death;
FW: cell survival pathway, intracellular signalling; AIDS; cancer; human; ss.
XX
XX
OS: Homo sapiens.
XX
PN: W09855507-A2.
XX
PD: 10-DEC-1998.
XX
PE: 01-JUN-1998; 98W01100255.
XX
PR: 11-SEP-1997; 9711.0121746.
PR: 05-JUN-1997; 9711.0121011.
PR: 30-JUN-1997; 9711.0121199.
XX
FA: (YEDA) YEDA RES & DEV CO LTD.
AA
PI: Boldin M, Malinin N, Wallach D;
XX
DE: W01.1999 97110121746
DE: P FDB: AA092795.
XX

[illegible]

Query Match	84.3%	Score 1769	DB 21	Length 2024	
Best Local Similarity	99.8%	Pred. No. 0			
Matches 1919	Conservative	0	Mismatches	3	Indels 0
			Gaps	0	

[illegible]

[illegible]

Query Match	78.8%	Score 1654	DB 21	Length 2709
Best Local Similarity	99.7%	Pred. No. 0		
Matches 1954	Conservative	0	Mismatches 6	Indels 0
			Claps 0	

[illegible]

Qy	1160	acattttcttcaagagctgtttat tcaagctaaagaaagaaagagtttcaagagctatttttcaaggtgcc	1219
Db	1167	acattttcttcaagagctgtttat tcaagctaaagaaagaaagagtttcaagagctatttttcaaggtgcc	1226
Qy	1220	atttcacctatgttgcacaaagaaagaaatggaatttatctctgcacatctacgtataaattcatggt	1279
Db	1227	atttcacctatgttgcacaaagaaagaaatggaatttatctctgcacatctacgtataaattcatggt	1286
Qy	1280	ccacaaagaggaattcatgttgcattccctctcagctcccatgaaatatagtgatttctccctgaaact	1333
Db	1287	ccacaaagaggaattcatgttgcattccctctcagctcccatgaaatatagtgatttctccctgaaact	1346
Qy	1340	tcaaggtccctcaccagctctctcaagacaaatgatttttttatctagaaaagctcgaagactgt	1399
Db	1347	tcaaggtccctcaccagctctctcaagacaaatgatttttttatctagaaaagctcgaagactgt	1406
Qy	1400	tatttttatgaaactgcattgcattctctgcgaatcagaatttgcattagaccacatttctgcga	1459
Db	1407	tatttttatgaaactgcattgcattctctgcgaatcagaatttgcgaayacaccacatttctgcga	1466
Qy	1460	tctcaaaaggaactgcatttctctgcatcacaagaaacattccatggtctcttcaggaataataaat	1519
Db	1467	tctcaaaaggaactgcatttctctgcatcacaagaaacattccatggtctcttcaggaataataaat	1526
Qy	1520	ccactctccactcgaaggaactcgaagagctctgcagagctagctatagccacagcagtggaac	1579
Db	1527	ccactctccactcgaaggaactcgaagagctctgcagagctggtatagccacagcagtggaac	1586
Qy	1580	cagagcaaaagaggaagacatttgttgaacaaatgacacaaagcctgccttaaccagctgccta	1639
Db	1587	cagagcaaaagaggaagacatttgttgaacaaatgacacaaagcctgccttaaccagctgccta	1646
Qy	1640	gattgcaccttctctccaggaacttgattcatgaaagaggaactatgaatttgttaagtaccaag	1699
Db	1647	gattgcaccttctctccaggaacttgattcatgaaagaggaactatgaacttggttaagtaccaag	1706
Qy	1700	ccctacaagggaactcgaaaatcagacaaattactagacactactgcacatccaaaggagaaqaa	1759
Db	1707	ccctacaagggaactcgaaaatcagacaaattactagacactactgcacatccaaaggagaaqaa	1766
Qy	1760	tttgcacaaagtatagtacaaaatttgaagatatacaaacagaatgggtcttcagaccttac	1819
Db	1767	tttgcacaaagtatagtacaaaatttgaagatatacaaacagaatgggtcttcagaccttac	1826
Qy	1820	ccgagaaataacttgttgtttctagatcaccatcttttaattttactctcaaaataaaaagcatg	1879
Db	1827	ccgagaaataacttgttgtttctagatcaccatcttttaattttactctcaaaataaaaagcatg	1886
Qy	1880	taagtgaacttgttttttcaagaaagaaatgtgttttcataaaaaggaatatttatatctctgttc	1939
Db	1887	taagtgaacttgttttttcaagaaagaaatgtgtgttctataaaaaggaatatttatatctctgttc	1946
Qy	1940	tttgactttttttatataaaattccgtgaattatttaaaactt	1979
Db	1947	tttgactttttttatataaaattccgtgaattatttaaaactt	1986

```

RESULT      4
AAK94554
ID          AAK94554 standard; cDNA; 2033 BP.
XX
AC          AAK94554;
XX
LT          06 NOV 2001 (first entry)
XX
DE          Human full-length cDNA, SEQ ID NO: 3453.
XX
RW          Human: full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
OS          Homo sapiens.
XX
PN          EP119094-A2.
XX

```

Qy	522	ctgaatttttttgggaattgaattactgaatagatggcaaatgaatcatgaatgaattcctac	591
Db	481	ctgaatttttttgggaattgaattactgaatagatggcaaatgaatcatgaatgaattcctac	549
Qy	582	atagqaaaaactgaatatcctgaattttacttggccatgaagatttcacatcctgaatgaaa	641
Db	541	atagqaaaaactgaatatcctgaattttacttggccatgaagatttcacatcctgaatgaaa	609
Qy	642	ttgccctttgaatgaatctactggcaaatatgaactcctcctttacttcattcatgaattga	701
Db	601	ttgccctttgaatgaatctactggcaaatatgaactcctcctttacttcattcatgaattga	669
Qy	702	aaqctcagaatattcttatggcaaatgaatttcattatgaagatttgcagattttgatttat	761
Db	661	aaqctcagaatattcttatggcaaatgaatttcattatgaagatttgcagattttgatttat	720
Qy	762	caaaagtggccgcatgatgtcctctcacagtcacaaagtacaaatctgcacacagaagaa	821
Db	721	caaaagtggccgcatgatgtcctctcacagtcacaaagtacaaatctgcacacagaagaa	789
Qy	822	ggcaaatattttatagcaactgaagaaatagcaactggcaaaaaatcaagggcagatga	881
Db	781	ggcaaatattttatagcaactgaagaaatagcaactggcaaaaaatcaaggggagatga	849
Qy	882	tcaagcacagatataatatacctatggcaatttatcacatgggaagtttatcgaagaaacag	941
Db	841	tcaagcacagatataatatacctatggcaatttatcacatgggaagtttatcgaagaaacag	909
Qy	942	cttttgaagattcaccaatctttttgcaaatgaattatgaatttatcagaaggaatcga	1001
Db	901	cttttgaagattcaccaatctttttgcaagatgaattatgaatttatcagaaggaatcga	969
Qy	1002	ctatttattaatgaagaaaatttggccatataatatacctcacccagacagatatactctc	1061
Db	961	ctgttatttaattgaagaaaatttggccatataatatacctcacccagacagatatactctc	1029
Qy	1062	taataagaaaattggaatgggcacaaaatccagatgaaagaacattctttcttaaatgtttaa	1121
Db	1021	taataagaaaattggaatgggcacaaaatccagatgaaagaacattctttcttaaatgtttaa	1089
Qy	1122	taaaacttgaacaaatttttgaagaaatttgaagagataaactttcttgaagctatttat	1181
Db	1081	taagacttgaacaaatttttgaagaaatttgaagagataaactttcttgaagctatttat	1149
Qy	1182	agctaaagaaaaaaagatttgcaggaattttgaagtgcaatttgaattatgaagaaagaa	1241
Db	1141	agctaaagaaaaaaagatttgcaggaattttgaagtgcaatttgaattatgaagaaagaa	1209
Qy	1242	aaatggaatttatctctgaacataacctgtaaatcatggtccaaagaaaggaatcatgtgga	1301
Db	1201	aaatggaatttatctctgaacataacctgtaaatcatggtccaaagaaaggaatcatgtgga	1269
Qy	1302	cctctcagctccatgaaaatagtggtttctctgaagcttcgaagctccctgcacgtcctc	1361
Db	1261	cctctcagctccatgaaaatagtggtttctctgaagcttcgaagctccctgcacgtcctc	1329
Qy	1362	aaqacaattgatttttttatctagaaaaagctcaagactgttattttatgaagctgaatcact	1421
Db	1321	aaqacaattgatttttttatctagaaaaagctcaagactgttattttatgaagctgaatcact	1389
Qy	1422	atcctgaaaatcacagtttcgaatagcaacattttctgaatttcgaagggctgaattctatga	1481
Db	1381	atcctgaaaatcacagtttcgaatagcaacattttctgaatttcgaagggctgaattctatga	1449
Qy	1482	atcacaaagacactccatgcttttcgaacaaatgaataacactctcaattgcaggaagat	1541
Db	1441	atcacaaagacactccatgcttttcgaacaaatgaataacactctcaattgcaggaagat	1509
Qy	1542	cagaagctctgcagactggaatatagccacagagatggaatccagagcaaaaaggaaggaattg	1601
Db	1501	cagaagctctgcagactggaatatagccacagagatggaatccagagcaaaaaggaaggaattg	1569
Qy	1602	gaagccaaatgacagaagctgaactgaagcaattcctgaagctcctttatccaggaatt	1661

```

RESULT 5
AAZ09246
ID: AAZ09246 standard, cDNA, 1931 BP.
XX
AC:
XX
AAZ09246:
DI: 25-OCT-1999 (first entry)
XX
DE: Human CARD-3 cDNA.
XX
KW: CARD-3; caspase recruitment domain; CARD-4; regulation; detection;
KW: caspase activation; detection; screening; therapy; diagnosis; disease;
KW: apoptotic cell death; Fas/AFK-1 receptor complex; TNF receptor complex;
KW: cancer; follicular lymphoma; carcinoma; p53 mutation; viral infection;
KW: hormone-dependent tumour; autoimmune disorder; Alzheimer's disease;
KW: systemic lupus erythematosus; immune-mediated glomerulonephritis; stroke;
KW: Parkinson's disease; amyotrophic lateral sclerosis; retinitis pigmentosa;
KW: spinal muscular dystrophy; cerebellar degeneration; anaemia; drug;
KW: myelodysplastic syndrome; myocardial infarction; cell proliferation;
KW: cell differentiation; cell survival; CARD-4L; CARD-4S; CARD-4Y;
KW: CARD-4Z; human; ds.
XX
OS: Homo sapiens.
XX
FH: Key Location/qualifiers
FI: CDS 214..1826
FT: /*tag: a
FI: /*product "CARD-3"
XX
PR: W9904182.1.
XX
PD: 12-AUG-1999.
XX
PF: 05-FEB-1999: 99WS-0802544.
XX
PR: 08-DEC-1998: 98US-0207459.
PR: 06-FEB-1998: 98US-0019942.
PR: 17-JUN-1998: 98US-0099041.
XX
PA: (MILL-) MILLENNIUM PHARM INC.
XX
PI: Bertin J.
XX
DR: WP1: 1999-494269/41.
DR: P-PSDB: AAY41140.

```

Query Match	77.1%	Score 1617	DB 26	Length 1941	
Best Local Similarity	59.7%	Prod. No. 0			
Matches 1867	Conservative	d	Mismatch 5	Indels 0	Gaps 0
ay	58	atgaagctatgaatttgaagaagaaagaggaataaataaaagcaatcgaaggaatgaggagccctc	117		
ab	12	atgaagctatgaatttgaagaagaaataaaggaatgagggaatcgaaggaatgaggagccctc	71		
ay	118	atgaagctatgaatttgaagaagaaaggaattttgaagaaatgaaggaatttgaaggggaatc	177		
ab	70	atgaagctatgaatttgaagaagaaaggaattttgaagaaatgaaggaatttgaaggggaatc	141		
ay	178	tgaaggaatgaagaaggaaggaatgaagaagcttgaaggaagacacgaaggaaggaaggaagga	237		
ab	132	tgaaggaatgaagaaggaaggaatgaagaagcttgaaggaagacacgaaggaaggaaggaagga	191		
ay	238	atgaagctatgaagaaggaaggaatgaagaaggaaggaatgaagaaggaaggaaggaaggaagga	297		
ab	132	atgaagctatgaagaaggaaggaatgaagaaggaaggaatgaagaaggaaggaaggaaggaagga	251		
ay	298	tgcctatgaagaaggaatgaagaaggaatgaagaaggaaggaaggaaggaaggaaggaagga	357		
ab	252	tgcctatgaagaaggaatgaagaaggaatgaagaaggaaggaaggaaggaaggaaggaagga	311		
ay	358	atgaagctatgaagaaggaaggaatgaagaaggaaggaatgaagaaggaaggaaggaagga	417		
ab	312	atgaagctatgaagaaggaaggaatgaagaaggaaggaatgaagaaggaaggaaggaagga	371		
ay	418	tgcctatgaagaaggaatgaagaaggaaggaatgaagaaggaaggaaggaaggaaggaagga	477		
ab	372	tgcctatgaagaaggaatgaagaaggaaggaatgaagaaggaaggaaggaaggaaggaagga	431		
ay	478	tgaatttgaattgaatttttgaatttttgaaggaatttgaattgaaggaatgaatttttgaagga	537		
ab	432	tgaatttgaattgaatttttgaatttttgaaggaatttgaattgaaggaatgaatttttgaagga	491		

```

597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

```

```

1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000

```

```

RESULT 6
AAF49001
ID AAF49001 standard; cDNA; 1931 bp.
XX
AC AAF49001;
XX
DI 23-APR-2001 (first entry)
XX
DE Human CARD; 3 cDNA.
XX
KW CARD-3; caspase recruitment domain; human; cancer; infection;
KW autoimmune disease; neurological disease; hematological disease;
KW immune disease; inflammation; antitumor; antiseptic;
KW immunomodulator; antiinflammatory; apoptosis; diagnosis;
KW gene therapy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 214..1826
FI Z=559 4
FI Znote "The open reading frame is also specifically
FT claimed in Claim 1(a)"
XX
PN W0200100826-A2.
XX
PD 04-JAN-2001.
XX
PE 28 JUN 2000; 2000W0 0817691.
XX
PR 28 JUN 1999; 99US 0340620.
XX
PA (MILL.) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WP1. 2001 061974/07.
DR P PSDB: AAB20079.
XX
PT Isolated intracellular proteins predicted to be involved in regulating
PI caspase activation are used for diagnosis and treatment of e.g. cancer,
PT viral infections, autoimmune diseases, neurological diseases and
PI hematological disorders
XX
PS Claim 1(a); Fig 1; 208pp; English.
XX
CC The present sequence is that of cDNA encoding human caspase
CC recruitment domain 3 (CARD-3; see AAB20079). The cDNA was isolated
CC following a database search using known CARD sequences. Plasmid

```

CC pXFL7A containing CARD-3 cDNA is deposited as ATCC 223037. CARD-3
 CC is an intracellular protein predicted to be involved in regulating
 CC caspase activation. It is useful as a modulating agent in
 CC regulating cellular processes include cell growth and cell death.
 CC Methods of diagnosing and treating patients suffering from a
 CC disorder associated with an abnormal level or rate of apoptotic
 CC cell death, abnormal activity of the Fas/CD95 receptor complex,
 CC abnormal activity of the tumour necrosis factor receptor complex
 CC or abnormal activity of a caspase involve administering a compound
 CC that modulates the expression or activity of CARD-3, CARD-4, CARD-5
 CC or CARD-6, a small molecule, antisense molecule, siRNA, ribozyme
 CC or polypeptide. Such disorders include cancer, viral infection,
 CC autoimmune disorders, neurological diseases, haematological
 CC disorders, inflammatory disorders and immune disorders. CARD
 CC nucleic acids can be used to express CARD proteins in a host cell
 CC e.g. for gene therapy applications, to detect a genetic lesion and
 CC to modulate CARD activity.
 CC
 XX
 SQ Sequence 1931 BP; 613 A; 428 C; 416 G; 474 T; 0 other.

Query Match 74.6%; Score 1566; DB 22; Length 1931;
 Best Local Similarity 99.7%; Prod. No. 0;
 Matches 1666; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 58 atcagctctggttcggaagaagcagcgctggcgtagggcatccgaggaatggcgccctc 117
 Db 12 atcagctctggttcggaagaagcagcgctggcgtagggcatccgaggaatggcgccctc 71
 Qy 118 atgacctgaatgattgaggaagcaaaaagggctcttgccggcctcgctcgttcagagggcgatc 177
 Db 72 atgacctgaatgattgaggaagcaaaaagggctcttgccggcctcgctcgttcagagggcgatc 131
 Qy 178 tggcgccctgagcgcgagcgtggaagccttgggaagccgcgcagcagagggcgacaccccgga 237
 Db 132 tggcgccctgagcgcgagcgtggaagccttgggaagccgcgcagcagagggcgacaccccgga 191
 Qy 238 accggcctgagcgcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 297
 Db 192 accggcctgagcgcccgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 251
 Qy 299 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 357
 Db 252 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 311
 Qy 358 gtcgcccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 417
 Db 312 gtcgcccgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 371
 Qy 418 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 477
 Db 372 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 431
 Qy 478 tgaatttgaatgacatttttcaaattttgggaatttgaatgagcctggaatttttgggaatt 537
 Db 432 tgaatttgaatgacatttttcaaattttgggaatttgaatgagcctggaatttttgggaatt 491
 Qy 538 agttactgaatgacattgcaaatggaatgaaatggaatggaatggaatggaatggaatggaatg 597
 Db 492 agttactgaatgacattgcaaatggaatgaaatggaatggaatggaatggaatggaatggaatg 551
 Qy 598 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 657
 Db 552 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 611
 Qy 658 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 717
 Db 612 tggcctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 671
 Qy 718 atggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 777
 Db 672 atggagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 731

Qy 778 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 837
 Db 732 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 791
 Qy 838 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 897
 Db 792 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 851
 Qy 898 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 957
 Db 852 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 911
 Qy 958 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1017
 Db 912 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 971
 Qy 1018 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1077
 Db 972 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1031
 Qy 1078 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1137
 Db 1032 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1091
 Qy 1138 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1197
 Db 1092 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1151
 Qy 1198 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1257
 Db 1252 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1211
 Qy 1258 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1317
 Db 1212 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1271
 Qy 1318 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1377
 Db 1272 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1331
 Qy 1378 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1437
 Db 1332 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1391
 Qy 1438 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1497
 Db 1492 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1451
 Qy 1498 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1557
 Db 1452 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1511
 Qy 1558 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1617
 Db 1512 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1571
 Qy 1618 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1677
 Db 1572 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1631
 Qy 1678 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1737
 Db 1632 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1691
 Qy 1738 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1797
 Db 1692 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1751
 Qy 1798 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1857
 Db 1752 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1811
 Qy 1858 atgctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1917

Query Match		72.4%	Score 518:	DB 21:	Length 2502:	
Best Local Similarity	99.6%	Prod. No. 0:				
Matches 1948:	Conservative	0:	Mismatches	6:	Indels	1: gaps
Qy	56	qqcaccacatctctctagaaagaaatctcagctctctgattctcqqaaqaaqaaqqcqqctcqqctcqqc	95			
Db	1	qqcaccacatctctctagaaagaaatctcagctctctgattctcqqaaqaaqaaqqcqqcctcqqctcqqc	60			
Qy	96	catccaaqqaatcqqcqqccctcatgaacctcattgctctcagaaqaaacaaagaaatctctaaqaa	150			
Db	61	catccaaqqaatcqqcqqccctcatgaacctcattgctctcagaaqaaacaaagaaqctcttgcctac	120			
Qy	156	ctcqqctctctcagaaqaaatctatctcagaaacccctgaagaaagaaatctcagaaacccctcagaaac	215			
Db	121	ctcqqctctctcagaaqaaqctctatctcagaaacccctgaagaaagaaatctcagaaacccctcagaaac	180			
Qy	216	cagaaqaaqaaqaaacacacccaaqaaacccqctcagaaqaaacccqaaacatgaacqaaqaaqaa	275			
Db	181	cagaaqaaqaaqaaacacacccaaqaaacccqctcagaaqaaacccqaaacatgaacqaaqaaqaa	240			
Qy	276	tcttqqcagcqqccctctgcctacacatctctctacacaaacctcqqcagaaacctcagcctacac	334			
Db	241	tcttqqcagcqqccctcagcctacacatctctctacacaaacctcqqcagaaacctcagcctacac	300			
Qy	335	cagaaqaaacccctctcagaaatctatctcatcccccctcagaaacaaactcagaaatctcagaaat	394			
Db	301	cagaaqaaacccctctcagaaatctatctcatcccccctcagaaacaaactcagaaatctcagaaat	360			
Qy	395	atgaagaaacctcagaaatccaaactccagctcctcagaaatcagaaacaaagaaatctcttgaag	454			
Db	361	atgaagaaacctcagaaatccaaactccagctcctcagaaatcagaaacaaagaaatctcttgaag	420			
Qy	455	qaagctcgaagattttacacaaagctcgaatttgaattacatttttccaaatttttqgaattttac	514			
Db	421	qaagctcgaagattttacacaaagctcgaatttgaattacatttttccaaatttttqgaattttac	480			
Qy	515	aattcagcctcgaatttttqgaattgaattactgaattacatgcgaattcgaattcatgaattga	574			
Db	481	aattcagcctcgaatttttqgaattgaattactgaattacatgcgaattcgaattcatgaattga	540			
Qy	575	ctcctacatcagaaacactgaattatctcagctctccttgcctcagaaatcgaattttcgaattcct	634			
Db	541	ctcctacatcagaaacactgaattatctcagctctccttgcctcagaaatcgaattttcgaattcct	600			
Qy	635	cattgaatttcgaccttcggtctgaatttacctcagaaatataaactcctcctttaccttcattcat	694			
Db	601	cattgaatttcgaccttcggtctgaatttacctcagaaatataaactcctcctttaccttcattcat	660			
Qy	695	gaacttcgaagactcgaagatattcttatcgaagaaatgaatttcattatcgaagtcgaagatttt	754			
Db	661	gaacttcgaagactcgaagatattcttatcgaagaaatgaatttcattatcgaagtcgaagatttt	720			
Qy	755	gaatttatcgaagtcgaagatgattcctctctcagaaatcagaaactcgaagaaatctcagaaac	814			
Db	721	gaatttatcgaagtcgaagatgattcctctctcagaaatcagaaactcgaagaaatctcagaaac	780			
Qy	815	qaagaaqaaqaaatctatttatatgcacctcgaagaaatctgaacccctgaagaaatcgaac	874			
Db	781	qaagaaqaaqaaatctatttatatgcacctcgaagaaatctgaacccctgaagaaatcgaac	840			
Qy	875	qccagatcagaaacagatattatagactatccaaattatcgaatcgaagaaatgattatccaa	934			
Db	841	qccagatcagaaacagatattatagactatccaaattatcgaatcgaagaaatgattatccaa	900			
Qy	935	aaagaaaccccttttgaagattctcagaaatctcttgcagaaatctatctgaattctcgaagaa	994			
Db	901	aaagaaaccccttttgaagattctcagaaatctcttgcagaaatctatctgaattctcgaagaa	960			
Qy	995	catcgaacctcttatgaatgaagaaatcttgcgaattatgattatccctcagcagaaagaaatata	1054			
Db	961	catcgaacctcttatgaatgaagaaatcttgcgaattatgattatccctcagcagaaagaaatata	1020			

Qy 1055 atctctctcaataaagaacttggatgagacacaaatctcagatgaagagccatctttcttcttaaa 1114
 Db 1021 atctctctcaataaagaacttggatgagacacaaatctcagatgaagagccatctttcttcttaaa 1080
 Qy 1115 tcttttaataaagaacttgaagacatttgaagagataacttttcttgaagct 1174
 Db 1081 tcttttaataaagaacttgaagacatttgaagagataacttttcttgaagct 1140
 Qy 1175 gttattcagctaaagaaaacaaagtacacagagctgttcaagtgccattccactatgtgac 1234
 Db 1141 gttattcagctaaagaaaacaaagtacacagagctgttcaagtgccattccactatgtgac 1200
 Qy 1225 aagaagaatagtgaatcttcttgaagatcctgaaatcattggtccacaaaggaatca 1294
 Db 1201 aagaagaatagtgaatcttcttgaagatcctgaaatcattggtccacaaaggaatca 1260
 Qy 1295 tatgaatcctctcagctccatgaagaatagtggttctctgaacatttcaagatccctgcc 1354
 Db 1261 tatgaatcctctcagctccatgaagaatagtggttctctgaacatttcaagatccctgcc 1320
 Qy 1455 gttctctcaagagcaatgattttttatctagaagaatctcaagagctgtttattatgaagctg 1414
 Db 1421 gttctctcaagagcaatgattttttatctagaagaatctcaagagctgtttattatgaagctg 1380
 Qy 1415 catcactctctctgaagaatcacaagtggagatagcaacatttctgagctcacaagagctgca 1474
 Db 1381 catcactctctctgaagaatcacaagtggagatagcaacatttctgagctcacaagagctgca 1440
 Qy 1475 ttctatgagctacaagagcaatcctcctatcttctgaagaaatcaaatcctcctcctcctga 1534
 Db 1441 ttctatgagctacaagagcaatcctcctatcttctgaagaaatcaaatcctcctcctcctga 1500
 Qy 1525 gaaacatcagaagctctggaagctgatatagctcaggaatggtccagagcagaagagaa 1594
 Db 1501 gaaacatcagaagctctggaagctgatatagctcaggaatggtccagagcagaagagaa 1560
 Qy 1595 gacatttgagaacaaatgacagaagcctgccttaccagtcctcctcctcctcctcctgctc 1654
 Db 1561 gacatttgagaacaaatgacagaagcctgccttaccagtcctcctcctcctcctcctgctc 1620
 Qy 1655 aaggaattgagctcagagaagagagctatgaacttctttagtaccagcctcacaagagcctca 1714
 Db 1621 aaggaattgagctcagagaagagagctatgaacttctttagtaccagcctcacaagagcctca 1680
 Qy 1715 aaagtccagcaatttaatagaacattatgacatcagaagaaagaatttgcacaaattata 1774
 Db 1681 aaagtccagcaatttaatagaacattatgacatcagaagaaagaatttgcacaaattata 1740
 Qy 1775 gtacaaaaattgaagaataacaaacaaatgggttttcaagccttaccacgaagaataacttgtg 1834
 Db 1741 gtacaaaaattgaagaataacaaacaaatgggttttcaagccttaccacgaagaataacttgtg 1800
 Qy 1835 atttctcagatcagcaatttcaatttacttcaatatagaagcattgaagtgaattgtttt 1894
 Db 1801 atttctcagatcagcaatttcaatttacttcaatatagaagcattgaagtgaattgtttt 1860
 Qy 1895 caagaagaagaatgttttctataaagaatatttatactcttctgttctgttgaattttttat 1954
 Db 1861 caagaagaagaatgttttctataaagaatatttatactcttctgttctgttgaattttttat 1920
 Qy 1955 ataaaaatccttgaattatgaagctt 1979
 Db 1921 ataaaaatccttgaattatgaagctt 1945

RESULT 8
 AAK92259
 ID AAK92259 standard; cDNA; 740 BP.
 XX
 AC AAK92259;
 XX
 DT 06-NOV-2001 (first entry)

XX
 DE Human cDNA 5' end sequence. SEQ ID NO: 719.
 XX
 KW Human; full length cDNA; cDNA synthesis; oligo capping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1130094-A2.
 XX
 PD 05 SEP-2001.
 XX
 FI 07 JUL 2000; 2000EP-0114089.
 XX
 PR 08 JUL-1999; 94JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02 MAY 2000; 2000JP-0184765.
 XX
 PA (HELI-) HELIX RES INSI.
 XX
 PI Ota T, Nishikawa T, Imai T, Hayashi K, Imeli S, Kawai Y,
 PI Wakamatsu A, Saito Y, Naito K, Kojima S, Otsuki T, Koga H.
 XX
 DR WPI; 2001-524255/58.
 XX
 PI 840 Primers useful for synthesizing full length cDNA clones and their
 PI use in genetic manipulation.
 XX
 PS Claim 2, SEQ ID No 719, 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5' and 3' ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is the nucleotide
 CC sequence of the 5' end of a cDNA provided in the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD ROM format directly from EPO.
 XX
 SQ Sequence 740 BP; 179 A; 193 C; 200 G; 164 T; 4 other;

Query Match: 22.4%; Score 468; DB 22; Length 740;
 Best Local Similarity: 99.6%; Prod. No. 2.6e-174;
 Matches: 568; Conservative: 0; Mismatches: 2; Indels: 0; Gaps: 0;

Qy 45 ctctagaaaataatgaagctctcttctcaagaaagagagagctgaagcctccagagc 194
 Db 1 ctctagaaaataatgaagctctcttctcaagaaagagagagctgaagcctccagagc 60
 Qy 105 aatgagagagct 164
 Db 61 aatgagagagct 120
 Qy 165 gaaagagagagct 224
 Db 121 gaaagagagagct 180
 Qy 225 gaaag 284
 Db 181 gaaag 240
 Qy 285 gaaag 344
 Db 241 gaaag 300
 Qy 445 ctgagcactatct 404
 Db 401 ctgagcactatct 360

Query Match: 99.99% Score: 498; 146 Zz Length: 745;
Identical Nucleotide Similarity: 99.99%; Prod. No.: 2,60-174;
Matches: 146; Mismatches: 2; Indels: 0; Gaps: 0;

```

RESULTS 10
AAF21841
ID AAF21841 standard; DNA; 463 BP.
XX
AC AAF21841;
XX
DT 27 MAR 2001 (first entry)
XX
DE Human breast and ovarian cancer associated antigen gene SEQ ID 228.
XX
KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neurotropic; neuroprotective; antiviral; antibacterial; hepatotropic;
KW antidiabetic; antitumour; antitubercular; antileukemic; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; diuretic; immune disorder;
KW Addison's disease; allergy; autoimmune haemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease; ds.
XX
OS Homo sapiens.
XX
PN W020005173 A1.
XX
PD 21 SEP 2000
XX
PF 08 MAR 2000; 2000W010508B1.
XX
PR 12 MAR 1999; 99B0124270.
XX
FA (HUMA ) HUMAN GENOME SOLI INC.
XX

```


PI Rosen CA, Ruben SM;
 XX
 DE WPI; 2:00-611515/58;
 DE P-PSDB; AAB58948;
 XX
 PT New human breast and ovarian cancer associated gene sequences and the
 PI polypeptides encoded by these genes, useful in the prevention,
 PI treatment and diagnosis of cancer, immune disorders, cardiovascular
 PI disorders and neurological diseases -
 XX
 PS Claim 1; Page 653; 129pp; English
 XX
 CC Sequences AAF22614 - AAF22631 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22042 - AAF22048 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic, immunosuppressive,
 CC neurotropic, neuroprotective, antiviral, antiallergic, hepatotropic,
 CC antidiabetic, antiinflammatory, anticancer, vulnerary, anticonvulsant,
 CC antibacterial, antifungal, antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis, cardiovascular disorders such as
 CC myocardial ischaemia; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy, and infectious diseases.
 XX
 SE Sequence 462 BP; 150 A; 92 C; 89 G; 131 T; 1 other.

Query Match: 16.7%; Score 35%; DE 21; Length 493;
 Best Local Similarity: 66.6%; Pred. No. 5 66-129;
 Matches 450; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 546 aatacatgccaaatggaatcattaaatgaactctacataggaagaaatgaatatcctatg 605
 Db 1 aatacatgccaaatggaatcattaaatgaactctacataggaagaaatgaatatcctatg 60
 Qy 606 tgaattggaattgagatttgaatcctgaatgaattgaccttgaattgaattacctga 665
 Db 61 tgaattggaattgagatttgaatcctgaatgaattgaccttgaattgaattacctga 120
 Qy 666 acaatataactcctccttcaattcatcatgaactgaagaactcagaatatcttattggaca 725
 Db 121 acaatataactcctccttcaattcatcatgaactgaagaactcagaatatcttattggaca 180
 Qy 726 atgaatttatggaattgagatttgaatttatcaattggaattgaattgaattgaattga 785
 Db 181 atgaatttatggaattgagatttgaatttatcaattggaattgagattgaattgaattga 240
 Qy 786 caaaattcagaatataagaaatctgaacacagatggaaggaacattatttatatgacactg 845
 Db 241 caaaattcagaatataagaaatctgaacacagatggaaggaacattatttatatgacactg 300
 Qy 846 aaaaactatgaacctggaacaaaaatcaagggtcagtatcaagcagcatatataatagctatg 905
 Db 301 aaaaactatgaacctggaacaaaaatcaagggtcagtatcaagcagcatatataatagctatg 360
 Qy 906 caattatcacatgggaattgttatccagaaaacagccttttgaagatatacacaatcctt 965
 Db 361 caattatcacatgggaattgttatccagaaaacagccttttgaagatatacacaatcctt 420
 Qy 966 tgcagataaatgtatagtatgtcacaaaggacat 997
 Db 421 tgcagataaatgtatagtatgtcacaaaggacat 452

RESULT 11

AAV13547
 ID AAV13547 standard; cDNA: 210 BP;
 XX
 AC AAV13547;
 XX
 DI 29 DEC 1998 (first entry)
 XX
 DE Clone 156108 cDNA fragment encoding a human protein kinase homolog.
 XX
 KW Protein kinase; cell signalling; inflammation; carcinoma; diabetes;
 KW human X linked adammaglobulinemia; nonspherocytic haemolytic anaemia;
 KW artherosclerosis; glioma; testisosis; cholera-based septic shock;
 KW 031064 Dictyostelium Y Kinase; ss.
 XX
 OS Homo sapiens.
 XX
 PN US5817479-A.
 XX
 PD 06-JUL 1998.
 XX
 PF 07-AUG-1996; 96US-0700575.
 XX
 PR 07-AUG-1996; 96US-0700575.
 XX
 PA (INVT) INVTTE PHARM INC.
 XX
 PI An Young J, Bandman G, Hawkins PR, Wilde CG;
 XX
 DE WPI, 1998 516307/47.

PI Human kinase polynucleotide(s) and recombinant products - useful
 PI for identification of modulators of the enzyme, and treatment of
 PI diseases associated with abnormal kinase expression
 XX
 PS Claim 1; Columns 45-48; 30pp; English.

CC The invention provides polynucleotides which encode novel protein
 CC kinase homologs expressed in various human cells and tissues. The
 CC present sequence represents the clone 156108 cDNA fragment derived from
 CC a human cDNA library designated THP-1 phorbol LPS. The cDNA encodes
 CC a protein kinase which shows homology to the AAO1064 Dictyostelium Y
 CC kinase. Vectors and host cells can be used for recombinant production
 CC of the protein kinase homolog. The recombinant proteins may be used to
 CC raise antibodies for use as anti-kinase therapeutics. Polynucleotides
 CC based on the polynucleotide sequences, i.e. probes and antisense
 CC constructs, the peptides and antibodies are claimed to be useful as
 CC tools for studying signalling cascades in cells and proteins, and for
 CC identifying inhibitors (drugs) to treat diseases and inflammatory
 CC conditions associated with abnormal kinase expression. Diseases that
 CC are claimed to be treatable include human X linked adammaglobulinemia,
 CC nonspherocytic haemolytic anaemia, artherosclerosis, carcinomas,
 CC diabetes, gliomas, testisosis, cholera-based septic shock, etc.
 XX
 SQ Sequence 210 BP; 74 A; 40 C; 41 G; 55 T; 0 other;

Query Match: 10.0%; Score 210; DE 19; Length 210;
 Best Local Similarity: 100.0%; Pred. No. 6 62-73;
 Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 844 tgaagaattatgaacttgaagaaaagaatcaagaaacagatatacagaacatataatagctatg 903
 Db 1 tgaagaattatgaacttgaagaaaagaatcaagaaacagatatacagaacatataatagctatg 60
 Qy 904 tgaattatcacatgggaattgttatccagaaaacagccttttgaagatatacacaatcctt 963
 Db 61 tgaattatcacatgggaattgttatccagaaaacagccttttgaagatatacacaatcctt 120
 Qy 964 tttcagataaatgtatagtatgtcacaaaggacatcaagccttttatataaagaagaatttt 1023
 Db 121 tttcagataaatgtatagtatgtcacaaaggacatcaagccttttatataaagaagaatttt 180
 Qy 1024 gccatataatataacctcaccgaacagatata 1053

```

RESULT 14
AAAF64725
ID   AAAF64725 standard; cDNA; 349 BP.
XX
AC   AAAF64725;
XX
DT   29-APR-2001 (first entry)

```

XX
DE Novel human polynucleotide, SEQ ID NO. 481.
XX
KW Human; cytostatic; gene therapy; colon cancer; prostate cancer;
KW breast cancer; lung cancer; cancer detection; ss.
XX
OS Homo sapiens.
XX
PN W0200102568-A2.
XX
PD 11 JAN 2001.
XX
PE 30 JUN 2000; 2000W010818374.
XX
PR 02 JUL 1999; 99US 0142310.
PR 02 JUL 1999; 99US 0142311.
XX
PA (CHIR) CHIRON CORP.
PA (HYSE) HYSEQ INC.
XX
PI Williams LR, Escobedo J, Innis MA, Garcia PD, Klinger J, Kassam A;
PI Reinhard C, Randazzo P, Kennedy CC, Est D, Lamson G, Drmanac R;
PI Orkenjakov R, Drmanac S, Dickson M, Lukat I, Leshkevitz D;
PI Kita D, Garcia V, Jones LW, Strachan-Grain B;
XX
DE WPI: 2001-091905/10.
XX
PT Library of polynucleotides for diagnosing a cancerous state of a
PT mammalian cell and detecting cancer, particularly of the colon or
PT prostate, comprises 3351 human polynucleotide sequences -
XX
PS Claim 9: Page 613; 1046pp; English.
XX
CC The present sequence is one of 3351 sequences in a library of human
CC polynucleotides. The library is used to detect differentially expressed
CC genes correlated with a cancerous state of a mammalian cell and can
CC detect colon, prostate, breast and lung cancer. The library can be used
CC to produce probes for detection of mRNA and to produce additional copies
CC of the polynucleotides. The probes can be used for chromosome mapping of
CC the polynucleotide and for detection of transcription levels. Ribozymes
CC or antisense oligonucleotides can be generated. The polynucleotides and
CC their gene products are used as genetic or biochemical markers (e.g. in
CC blood or tissues) that will detect the earliest changes along the
CC carcinogenesis pathway and/or monitor the efficacy of therapies and
CC preventive interventions. The polynucleotides, polypeptides and
CC antibodies against them can be used in pharmaceutical compositions to
CC treat the cancers and proliferative disorders such as neoplasia,
CC dysplasia and hyperplasia.
XX
SQ Sequence 349 BP; 126 A; 70 C; 58 G; 95 T; 0 others;

Query Match 1.89% Score 38; DB 22; Length 449;
Best Local Similarity 100.0%; Pred. No. 26 05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2061 aattttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
Db 247 aattttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 284

RESULT 15
AAC79976
ID AAC79976 standard; cDNA; 2082 BP.
XX
AC AA:79976;
XX
DI 12 FEB 2001 (first entry)
XX
DE Human secreted protein encoding cDNA for gene 29.
XX
KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;

KW neotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; vulnerary; gene therapy; treatment; autoimmune disease;
KW hyperproliferative disorder; cardiovascular disorder; ocular disorder;
KW cerebrovascular disorder; nervous system disorder; infectious; skin aging;
KW wound healing; epithelial cell proliferation; transplantation; ss.
XX
OS Homo sapiens.
XX
PN W0200058357-A1.
XX
PD 05 OCT 2000.
XX
PE 24 MAR 2000; 2000W010807223.
XX
PR 26 MAR 1999; 99US 1126506.
PR 07 JAN 2000; 2000US 1174852.
XX
PA (HUMA) HUMAN GENOME SCI INC.
XX
PI Ruben MA, Ruben SM, Komatsuilis G;
XX
DE WPI: 2000-611704/56.
DE P-PDB: AAB45053.
XX
PI Nucleic acid molecules encoding human secreted proteins, used in
PI preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PI Parkinson's diseases and cancers -
XX
PS Claim 1a: Page 356-357; 418pp; English.
XX
CC This invention describes novel isolated nucleic acid molecules (I)
CC encoding a human secreted proteins (II) which have immunosuppressive,
CC antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant,
CC vasotropic, cerebroprotective, neotropic, neuroprotective, antibacterial,
CC virucide, fungicide, ophthalmological and vulnerary activity, and as be
CC used for gene therapy. (I) and (II) are used to prevent, treat or
CC ameliorate a medical condition in e.g. humans, mice, rabbits, goats,
CC horses, cats, dogs, chickens or sheep. (I) and (II) are also used in
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition. The antibodies to (II) can also be used in alleviating
CC symptoms associated with the disorders and in diagnostic immunoassays
CC e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA).
CC Disorders which are diagnosed or treated include autoimmune diseases
CC e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms
CC of the breast or liver, cardiovascular disorders e.g. cardiac arrest,
CC cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous
CC system disorders e.g. Alzheimer's disease, infections caused by bacteria,
CC viruses and fungi and ocular disorders e.g. corneal infection. The
CC polypeptides can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, to support cell culture of primary tissues,
CC to regenerate tissues and in chemotherapy. The polypeptides can also be
CC used as a food additive or preservative to increase or decrease storage
CC capabilities.
XX
SQ Sequence 2082 BP; 568 A; 460 C; 423 G; 611 T; 0 others;

Query Match 1.89% Score 38; DB 21; Length 2082;
Best Local Similarity 100.0%; Pred. No. 1,6e+05;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2061 aattttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
Db 2029 aattttgaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2056

Search completed: May 14, 2002, 21:31:01
Job time: 8577 sec


```

RESULT 1
AB2777
LOCUS AB2777 2098 bp DNA Linear PAT 21-JAN-2000
DEFINITION Sequence 2 from Patent W09855507.
ACCESSION AB2777
VERSION AB2777.1 11:67s2464
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
unclassified.
REFERENCE
1 (bases 1 to 2098)
AUTHORS Boldin, M., and Wallach, D.
TITLE MODULATORS OF INTRACELLULAR INFLAMMATORY CELL DEATH AND CELL
SURVIVAL PATHWAYS
JOURNAL Patent, W 0855507 A 2 16 DEC 1998;
BOLDIN MARK (11); WALLACH DAVID (11)
FEATURES
Location/Qualifiers
Source 1..2098
/organism "unidentified"
/db_xref "taxon:32644"
BASE COUNT 649 a 452 c 449 g 539 t
ORIGIN

```

QY	1021	TTTGGCATATGATATACCTGAGGAGGAGGATATGATCTCTTAAATGAAAGATGGATGGG	1088
DB	1021	TTTGGCATATGATATACCTGAGGAGGAGGATATGATCTCTTAAATGAAAGATGGATGGG	1088
QY	1081	ACAAAATCCAGATGAAAGAGCTATCTTCTTAAATGTTTAAATAGAGCTTGAAGCACTTTT	1140
DB	1081	ACAAAATCCAGATGAAAGAGCTATCTTCTTAAATGTTTAAATAGAGCTTGAAGCACTTTT	1140
QY	1141	GAGAACTTTGAAGAGATAACTTTCTTGAAGCTGTTATCTAGCTTAAAGAAAACAAAGCTT	1200
DB	1141	GAGAACTTTGAAGAGATAACTTTCTTGAAGCTGTTATCTAGCTTAAAGAAAACAAAGCTT	1200
QY	1201	ACAGAGCTGTTTCAAGTGGCATTTCAGCTATGTGTGCAAGAAGAAAATGGAAATATCTCTTGAA	1260
DB	1201	ACAGAGCTGTTTCAAGTGGCATTTCAGCTATGTGTGCAAGAAGAAAATGGAAATATCTCTTGAA	1260
QY	1261	CATACCTGTAAATCATGCTGCCACAAAGAGGAACTATGTGGAATGCTCTTACGCTCCATGAAAA	1320
DB	1261	CATACCTGTAAATCATGCTGCCACAAAGAGGAACTATGTGGAATGCTCTTACGCTCCATGAAAA	1320
QY	1321	TAGTGGTTCTGCTGAAAGCTTCAAGCTGCTTGGCAGCTCTCTAAGCAAAATGATTTTATACT	1380
DB	1321	TAGTGGTTCTGCTGAAAGCTTCAAGCTGCTTGGCAGCTCTCTAAGCAAAATGATTTTATACT	1380
QY	1381	TAGAAAGCTTCAAGAGCTGTATTTATGAAAGTGTGATCTATGCTCTGAAATTCAGAGCTT	1440
DB	1381	TAGAAAGCTTCAAGAGCTGTATTTATGAAAGTGTGATCTATGCTCTGAAATTCAGAGCTT	1440
QY	1441	GGATAGCACTTCTTGGATCTCAAAAGGCTGTGATCTCTGAACTCAAGAGCACCTCTATG	1500
DB	1441	GGATAGCACTTCTTGGATCTCAAAAGGCTGTGATCTCTGAACTCAAGAGCACCTCTATG	1500
QY	1501	CTCTTCAGCAATAATAAATGCATCTCTCAACTGCAAGGAACTCAGAAAGCTCTGCAAGCTGG	1560
DB	1501	CTCTTCAGCAATAATAAATGCATCTCTCAACTGCAAGGAACTCAGAAAGCTCTGCAAGCTGG	1560
QY	1561	TATAGCTCAGCAGTATCTCAGAGCAAAAGGGGAAGACATGTGGAACAAATGACAGAACT	1620
DB	1561	TATAGCTCAGCAGTATCTCAGAGCAAAAGGGGAAGACATGTGGAACAAATGACAGAACT	1620
QY	1621	CTGCTTAAACAGTCTGATGCTCTCTGCTCAAGGATGATCATGAAAGAGGACTA	1680
DB	1621	CTGCTTAAACAGTCTGATGCTCTCTGCTCAAGGATGATCATGAAAGAGGACTA	1680
QY	1681	TGAAGCTGTGATGATCAAGGCTACAAAGGACTCAAAAGTACAGCAATTAAGTACAGACTA	1740
DB	1681	TGAAGCTGTGATGATCAAGGCTACAAAGGACTCAAAAGTACAGCAATTAAGTACAGACTA	1740
QY	1741	TGACATCTCAAGGACAAAGATTTGCCAAAGTTATAGTACAAAAATGAAAGATAACAAACA	1800
DB	1741	TGACATCTCAAGGACAAAGATTTGCCAAAGTTATAGTACAAAAATGAAAGATAACAAACA	1800
QY	1801	AATGGCTCTTCAGCTCTTACCTCGGAAATAGCTGTGGTTCTAGATCAACATCTTTAAATTT	1860
DB	1801	AATGGCTCTTCAGCTCTTACCTCGGAAATAGCTGTGGTTCTAGATCAACATCTTTAAATTT	1860
QY	1861	ACTTCAAAATAAAAAGCACTAAGTCACTGTTTTCAAGAGAAAATGCTTTTCATAAAAAG	1920
DB	1861	ACTTCAAAATAAAAAGCACTAAGTCACTGTTTTCAAGAGAAAATGCTTTTCATAAAAAG	1920
QY	1921	ATATTTATATCTCTGCTGCTTTCAGCTTTTATATATAAAATCTGATTAAGCTTCT	1980
DB	1921	ATATTTATATCTCTGCTGCTTTCAGCTTTTATATATAAAATCTGATTAAGCTTCT	1980
QY	1981	AWWRAAKIKTCTTTSKTAATATTAQTCTCTCTCATGAGCTGAGATTTTTTTTAA	2040
DB	1981	AWWRAAKIKTCTTTSKTAATATTAQTCTCTCTCATGAGCTGAGATTTTTTTTAA	2040
QY	2041	TTAAATCAAAATGAAAGTTTGAATTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	2098
DB	2041	TTAAATCAAAATGAAAGTTTGAATTTGAAAAAAGAAAAAAGAAAAAAGAAAAAAG	2098

Query Match 81.9%; Score 1719; DB 9; Length 1898;

[illegible]

```

136 1190 GATACAAAAATACAGATGAAGAGATATCTTTTAAAAATGTTAATAGAACTTGAAATAGT 1189
137 1191 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1197
138 1192 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1198
139 1193 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1199
140 1194 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1200
141 1195 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1201
142 1196 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1202
143 1197 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1203
144 1198 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1204
145 1199 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1205
146 1200 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1206
147 1201 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1207
148 1202 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1208
149 1203 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1209
150 1204 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1210
151 1205 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1211
152 1206 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1212
153 1207 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1213
154 1208 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1214
155 1209 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1215
156 1210 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1216
157 1211 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1217
158 1212 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1218
159 1213 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1219
160 1214 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1220
161 1215 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1221
162 1216 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1222
163 1217 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1223
164 1218 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1224
165 1219 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1225
166 1220 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1226
167 1221 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1227
168 1222 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1228
169 1223 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1229
170 1224 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1230
171 1225 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1231
172 1226 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1232
173 1227 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1233
174 1228 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1234
175 1229 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1235
176 1230 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1236
177 1231 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1237
178 1232 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1238
179 1233 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1239
180 1234 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1240
181 1235 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1241
182 1236 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1242
183 1237 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1243
184 1238 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1244
185 1239 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1245
186 1240 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1246
187 1241 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1247
188 1242 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1248
189 1243 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1249
190 1244 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1250
191 1245 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1251
192 1246 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1252
193 1247 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1253
194 1248 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1254
195 1249 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1255
196 1250 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1256
197 1251 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1257
198 1252 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1258
199 1253 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1259
200 1254 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1260

```

```

RESULT 1 4
AX092101
1.00US AX092101 1941 bp DNA Linear PAT 27 FEB 2001
DEFINITION Sequence 1 from Patent WO0100826.
ACCESSION AX092101
VERSION AX092101.1 GI:14170986
FEATURES
ORIGIN
ORGANISM Homo sapiens
Taxonomy: Metazoa; Chordata; Mammalia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE
AUTHOR (Baron R et al 1991)
TITLE A gene, antibodies of the gene-related protein family and uses thereof

```

```

JOURNAL Patent: WO 0100826 A 1 04 JAN 2001;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1.1931
organism "Homo sapiens"
DB_xref "taxon:9606"
214.1.1836
CDS
/note "unnamed protein product"
/codon_start 1
/protein_id "CAC34155.1"
/db_xref "GI:14170986"
/translation "MNGEATC SALETTFYHKLADLRKYL SRGASCTVSSAHADWRVAV
AVKHLHILHLLSEKKEVLEKAFLEHKAESLLEPLDGLNEFEELAGVLDMPNGS
LNR LHRKLEHYDVAVLEKRLLEHLEAGVNYLHNMIPPLHHLHLEKQNLLENEFHV
KLADPGLSKWRKMSLSGRSSSKSAFEGGLTYMILENYELGOKSKASTKHGLTSYAVI
TWEVLSRKQPFEDVTNLEQIMYSVSQGBRPVINEESLPYDIPHRARMSLEIESGWAQN
DEKESFLKLTLEKPVLETFELTLEAVLQKKELEQSVSSATHELEKAMELSLN
LIVNHGQURESCGSSQLHENSQSPETSKSLPAQDNDFLSRKAQDQYFMKLHHCQGNH
SWISTLSGSRAAFCHKLTFCSSALINLTAGNSPKQDGTAGQWLEKREKLVNQ
MFEACTNQSLHALLSKRLIMKEEYEEVSEKLEHSEVQLLEHLEKQEFKAVIVLEK
LEKLRKMLDQYFELIVVSRKSLNLLQKSM"
BASE COUNT 613 A 430 C 416 G 472 T
ORIGIN

```

```

Query Match: 79.5%; Score 1668; DB 6; Length 1941;
Best Local Similarity: 99.8%; Pred. No. 0;
Matches 1868; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

136 1190 GATACAAAAATACAGATGAAGAGATATCTTTTAAAAATGTTAATAGAACTTGAAATAGT 1189
137 1191 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1197
138 1192 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1198
139 1193 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1199
140 1194 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1200
141 1195 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1201
142 1196 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1202
143 1197 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1203
144 1198 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1204
145 1199 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1205
146 1200 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1206
147 1201 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1207
148 1202 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1208
149 1203 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1209
150 1204 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1210
151 1205 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1211
152 1206 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1212
153 1207 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1213
154 1208 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1214
155 1209 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1215
156 1210 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1216
157 1211 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1217
158 1212 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1218
159 1213 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1219
160 1214 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1220
161 1215 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1221
162 1216 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1222
163 1217 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1223
164 1218 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1224
165 1219 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1225
166 1220 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1226
167 1221 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1227
168 1222 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1228
169 1223 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1229
170 1224 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1230
171 1225 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1231
172 1226 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1232
173 1227 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1233
174 1228 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1234
175 1229 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1235
176 1230 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1236
177 1231 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1237
178 1232 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1238
179 1233 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1239
180 1234 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1240
181 1235 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1241
182 1236 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1242
183 1237 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1243
184 1238 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1244
185 1239 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1245
186 1240 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1246
187 1241 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1247
188 1242 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1248
189 1243 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1249
190 1244 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1250
191 1245 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1251
192 1246 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1252
193 1247 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1253
194 1248 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1254
195 1249 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1255
196 1250 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1256
197 1251 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1257
198 1252 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1258
199 1253 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1259
200 1254 TTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAAAATTTTAAA 1260

```


Qy 718 atttggacaatgaatttccatgtttaagattgcagatttgggtttatcaaagtggcgcatgat 777
Db 672 ATTGGACAATGAATTTTCATGTTAAAGATGCGAGATT TGGTTTATCAAAGTGGCGCATGAT 731
Qy 778 qtccctctccacagtcacgaagtacgaatctgcac tagaaggaggggacaattatttata 837
Db 732 GTCCCTCTCACAGTCACGAAGTAGCAAAATCTGCAC TAGAAGGAGGGACAATTATCTATAT 791
Qy 838 gccacctgaaaactatgaacctggacaaaaaatcaa iggccagtatcaagcacgatata 897
Db 742 GCCACCTGAAAACATAACAACCTGGACAAAAATCAA IGGCCAGTATCAAGCACGATATATA 851
Qy 898 tagctatcagtttatcacatgggaagtgttatccajaaaacagcccttttgaagatgtcac 957
Db 852 TAGCTATCAGTTTATCACATGCCAAGCTTTTATCCA AAAACAGCCCTTTTGAAGATGTCA 911
Qy 958 caatcccttgccagataatgttatagtgatcacaaagacatcgacctgttattatgaaga 1017
Db 912 CAATCCCTTGCGAGATAATGTATAGTGTGTCAAAAGACATCGACCTGTATTAATGAAGA 971
Qy 1018 aagttttgcatatagatataacctccagagagcaagtgatctctcttaatatagaagtgagtg 1077
Db 972 AACCTTGCCATATGATATACCTTCACGACGACGCAACATCGCTCTTAATACAAAGTGCATC 1031
Qy 1078 ggcacaaaatccagatgaagagccatcttttcttaaatgttttattagaacttgaacagtg 1137
Db 1032 GGCACAAAATCCAGATGAAGAGCCATCTTTCTTAATATGTCTTAATACAACTGAACCAAG 1091
Qy 1138 tttagagacatttgagagatgaacttttcttgaagtgtttatttcagctagaagagacaaa 1197
Db 1092 TTTAGAACATTTGAAGAGATAACTTTTCTTGAAGTGTCTTATCGCTTAAGAAAAACAAA 1151
Qy 1198 gtttccagagtggtttccagtgccattccctatgtgacaaagaaatggatttatctct 1257
Db 1152 GTTACAGAGTGTTTCAAGTGCCTATTCACCTATGTGCAAGAAAGAAATGGAAATTATCTCT 1211
Qy 1258 ggcacataatgtgaactgaagtgacaaagagggaatcatgtgggtctctcagctccatga 1317
Db 1212 GGCACATAATGTGAATGAGCTGACAAAGAGCAAAATATCTGATCTCTCAACCTCATGA 1271
Qy 1318 aaatagtggtttctccgtaaaacttcaagggtccctgcagctccctcaagacaatgatttttt 1377
Db 1272 AAATAGTGGTTCTCTCGTAAACTTCAAGCTCCCTGCTAGCTCCCTCAAGACAAATGATTTTT 1331
Qy 1378 atctagaaaaagctcaagactgttatttttgaagcggcatcaactgtctctggaatcacag 1437
Db 1332 ATCTAGAAAAAGCTCAAGACTGTATTTTATGAAGCTGCAATCACTGCTGGAAATCACAG 1391
Qy 1438 ttggatagacacattctcagatctcagaagggttgattctgtgatacaagaacacac 1497
Db 1392 TTGGATAGACACATTCTCAGATCTCAAAAGGCTCTATCTCTGATCACAAGACCATCT 1451
Qy 1498 atgtctttgaagaaatgaatgaactctcaactgttaggaatcagaaggtcttcagagc 1557
Db 1452 ATGTCTTTAGAAATGAATGAATCTCAACTCTCAACTCTAGGAAATCAGAAGCTCTCAGAGC 1511
Qy 1558 tgggtatgagcagagtgagtgatcagaagcaaaatgttagacatttgtaacaaatgacaga 1617
Db 1512 TGGGTATGAGCAGAGTGAAGTGAATGAAGCAAAAGTGAAGCAATCTCAAAATGAAGAGA 1571
Qy 1618 aggttgatgaagagtgatgagtgatgagtgatgagtgatgagtgatgagtgatgagtgag 1677
Db 1572 AGGTGATGAAGAGTGAAGTGAATGAAGTGAATGAAGTGAATGAAGTGAATGAAGTGAATGA 1631
Qy 1678 ctatgaatttggttagtgaagagcttagaggaacttagaagtcagacaattactagacac 1737
Db 1632 CTATGAATTGTTTAGTGAAGAGCTTAGAGGACCTTAGAAGTCAGACAATTACTAGACAC 1691
Qy 1738 tactgacatcagaaggaggaaggaattttgcgaagttaagtcacaaaattgaaagataacaa 1797
Db 1692 TACTGACATCAAGGAGGAAGGAATTTGCCAAGTTATAGTACAAAAATTGAAAGATAACAA 1751
Qy 1798 acagaatgggtcttcagcccttaccgcgaataacttgagggtttctagatcagccattttaa 1857

Db 1752 ACAAATGGGTCTTCAGCCCTACCCGGAATACCTGTGGTTCTAGATCACCATCTTTAAA 1811
Qy 1858 tttaacttcaaaaataaaagatgtaagtgaactgttttttaagaagaagtgtgtttcataaa 1917
Db 1812 TTTACTTCAAAAATAAAGCATGAAGTGACTGTTTTTCAAGAAGAAATGTGTTTCATAAA 1871
Qy 1918 agqatatttata 1929
Db 1872 AGGATATTATA 1884

RESULT 4

AF027706

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

PEPATTRES

source

gene

CDS

polyA_signal

polyA_site

polyA_signal

polyA_site

BASE COUNT

ORIGIN

1501 bp mRNA linear PPI 10-SEP-1998

Homo sapiens serine/threonine kinase RICK (RICK) mRNA, complete cds.

AF027706.1 GI:3123886

human.

Homo sapiens

1 (bases 1 to 2501)

Inohara,N., del Peso,L., Koseki,T., Chen,S. and Nunez,G.

RICK, a novel protein kinase containing a caspase recruitment domain, interacts with CLARP and regulates CD95-mediated apoptosis

J. Biol. Chem. 273 (26), 12266-12268 (1998)

98241596

2 (bases 1 to 2501)

Inohara,N., Koseki,T., Chen,S., del Peso,L. and Nunez,G.

Direct Submission

Submitted (01-OCT-1997) Dept. Pathology, Comprehensive Cancer & Geriatrics Center, 4-131 CCGC 1500 E. Medical Center Dr, Ann Arbor, MI 48109, USA

Location/Qualifiers

1. 2501

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="8"

/map="8q21"

1. 2501

/gene="RICK"

25. 1847

/gene="RICK"

/function="regulates apoptosis"

/note="CARD domain protein; phosphorylated protein; phosphorylated protein; apoptosis regulator"

/codon_start=1

/product="serine/threonine kinase RICK"

/protein_id="AAC34970.1"

/db_xref="GI:3123887"

/translation="MNGEATCSALPTIPYHKIADLEYLSKAGSTIVSSARHADWRVQVAVFLHGGTLLDSEKLVIEATDIFKAFVYIIEHITCNFFFFCTVTEYMPNGSLNELLRKTEYPDVAWPLRPRILHEIALGVNYLHNMTPELLHDLKTONILLDNFPHVLAAGFLKAWMMELLSKSSSSAEGCELYMFTENYEPQKSFASIAHDYSYAVITWEVLSRKQPHEDVINPLOTIMYSVSGHRPVINKEESLPYDIPHRARMSLIESGWAQNDKRRSEIKVLEFVIRKPRKPRKRAVIRKKIKRVSSAHLCKKKKPISTNIPVNHGQKESQSSQLHENSQSPETSRSLDAPQNDPISRKACDQYFMKLHHCQGNHWSSTLSGQRAAPCRKKTTPCSSATINPISTAGNSPKLPQTAQQWIQSKREDIVNQHTEACLNQSLCALLSPDLIMKEDYELVSTKPTSKVPQIDTTPICQEFKAVIVQKLEPKQKQIGVYVFTIVVSFSPSTNLIKNSM"

1833. 1838

/gene="RICK"

1869

/gene="RICK"

2479. 2484

/gene="RICK"

2501

/gene="RICK"

768 a 535 c 499 q 699 t

```

RESULT 5
AF064824
LOCUS       AF064824                1902 bp    mRNA    linear    PRI 07 JUN 1998
DEFINITION  Homo sapiens CARD-containing ICE associated kinase mRNA, complete cds.
ACCESSION   AF064824
VERSION     AF064824.1    GI:429017

```

[illegible]

Qy	192	cgctctgacacggggcgtggagcctctggagggcgccggcggggggacacacgggaaacc	241
Db	165	GGGCTTGAAGGGGGGGTGGCAAGCTTCTGGGAGGCGGCTTCAGATAGGGGGGCACACGGCGAAACCG	224
Qy	242	gcctgagcggccgggggacatgaacggggaggcgcaatctgcagcgccctggccacacatttcc	301
Db	225	GGGCGAGCGGGCGCGGACATGAACGGGGTAGGGCAATTCGCAAGCGGCTCTGGCCACCAATTCGCG	284
Qy	302	tacacaaaaactctgcggaccttagctacctgagccgctggcgcctctggcacttgatctgctcc	361
Db	285	TACCAACAAACTCTGCGCACTTGGCTAAGTGAAGAGCTGGAGAGGCTCTGGCACTGTGTGCTGC	344
Qy	362	gcgcgcacacgcagacactgagcgcgctcgaagtggcgctgaagcacctgcacatccacacctcc	421
Db	345	GCGCGCCACGCGACACTGGCGGCTCTAGGTGGCGCTTAAGTAACTGACATCCACACCTCGG	404
Qy	422	ctactgcgcagctgaaagaaagagatgttttaagagcagctgaaattttacacaaagctaga	481

[illegible]

```

21  00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
22  00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
23  00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000

```

```

FEATURES             1 to 600
ACCESSION            AF078590
LENGTH               1623 bp      mRNA       linear       PKI 28 JUL 1998
DEFINITION           Homo sapiens receptor interacting protein 2 (RIP2) mRNA, complete cds.
AUTHOR              Aikawa M,
VERSION             AF078590.1 GI:5442669
KEYWORDS             Human;
                    Homo sapiens;
                    Eukaryota; Metazoa; Chordata; Chnaria; Vertebrata; Euteleostomi;
                    Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE            1. (bases 1 to 1623)
AUTHORS              Moriarty J.V., Nijl, and Dixit V.M.
TITLES              RIP2 is a novel NF-kappaB activating and cell death inducing kinase
JOURNAL              Blood, Chem. 273 (27), 16968-16975 (1998)
MISC_FEATURE         cDNA clone.
REFERENCE            1. (bases 1 to 1623)
AUTHORS              Moriarty J.V., Nijl, and Dixit V.M.
TITLES              RIP2 SubMISSION
JOURNAL              Submitted (15 JUN 1998) Molecular cloning, Genentech Inc., 1 DNA
                    Way, South San Francisco, CA 94065, USA
ORIGIN              1 position/qualifiers
                     1..1623
                     /translation="Homo sapiens"
                     /db_xref="taxon:9606"
                     1..1623
                     /gene="RIP2"
                     1..1623
                     /name="Rip2"
                     /function="activates NF-kappaB"
                     /function="induces cell death"
                     /note="RIP2: serine/threonine kinase homolog, contains
                     "Akt motif"
                     /sdon_start=1
                     /product="receptor-interacting protein 2"
                     /protein_id="AAC27722.1"
                     /db_xref="GI:5442669"
                     /translation="MNCEATCSALFFTPYHKLAQLKYLSEGGASGVSSARHADWRVQV
                     AVKLRLHTPLDSEPEVLPPFAELHKKARESYLLPGICNEPEFLGIVTEYPNGS
                     INEHLHKLEPYPAWAPLGRKILAEIALGVNYLRNMFTPLTHLDLKQTNILLNEEFIV
                     KADGLSKRWMMSSSSKSSKSAAGVLIYYMRISPEYDQSKSKASKHDLSYAVI
                     LWEVLRKKQPFEIVLNELQIMYSVSQHPRVINEESTLPYDIPHRMTSLIESCWAQN
                     PLERFSETKELLKEGFVERLEEELFLRAVLQLEKETKLQSVSVAHQTEKKMLSLN
                     LPVNHGDEESQGSSQCHENSQSFETSRSLPAPQDNDFLSRKAAQCYFMKLHPCGNH
                     QWGTLDSGSQAAPLGAKTIPTSSSATINLETTAINSIKLAGSTAGQWQSGKRIIVNQ
                     MEATLNQSLDALSLKSLLMKRELYELSVSLKELTSKVQRDLDTIQWEEKAAVTVQ

```

```

RESULT 7
AX082201
LOCUS          AX082201                1620 bp            DNA            linear            PAT 27-FEB-2001
DEFINITION     3' UTR of a human Patent WO/980826.
ACCESSION      AX082201
VERSION        AX082201.1  GI:12179988
KEYWORDS
SOURCE
   ORGANISM     Homo sapiens
                Eukaryota; Metazoa; Chordata; Mammalia; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1620)
AUTHORS       Bertin, J.
TITLE          Novel molecules of the card related protein family and uses thereof
JOURNAL        Patent: WO 980826-A 3 04 JAN 2001;
                Millennium Pharmaceuticals, Inc. (US)
FEATURES
   source       Location/Qualifiers
                1..1620
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT     517 a      360 c      320 g      423 t
ORIGIN

```

[illegible]


```
repeat_region 34251..34949  
/rpt_family: "MatK"  
34999..35256  
/rpt_family: "Alu"  
35384..35477  
/rpt_family: "MEK2-type"  
35941..36286  
/rpt_family: "Alu"  
36686..36985  
/rpt_family: "Alu"  
37488..37673  
/rpt_family: "L2"  
37862..38161  
/rpt_family: "Alu"  
38297..38599  
/rpt_family: "Alu"  
41954..42006  
/rpt_family: "L2"  
42596..42777  
/rpt_family: "MEK1-type"  
43357..43658  
/rpt_family: "Alu"  
44088..44386
```

Cy	1541	ttgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1609
Dh	49877	ttgagaaatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	49936
Cy	1661	gtgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1690
Dh	49977	gtgagaaatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	49999
Cy	1681	ttgagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1720
Dh	49937	ttgagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	40056
Cy	1721	agagaaatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1780
Dh	40057	agagaaatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	40116
Cy	1791	aatatgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1840
Dh	40117	aatatgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	40176
Cy	1941	aatatgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1900
Dh	40117	aatatgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	40236
Cy	1961	gaagatctctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1960
Dh	40237	gaagatctctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	40296
Cy	1961	tcggtgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	1979
Dh	40297	tcggtgaagatctctctgaaatctggtatgagctgaggaatcatctcgaagcgaaggaaggaatatt	40415

```

RESU11  9
AF117829          220250 bp    DNA      linear    PRI 13 JAN 1999
LOCUS      Homo sapiens Hsp111, PINK gene, complete, 5' capped.
DEFINITION
ACCESSION  AF117829
VERSION    AF117829.1  GI:14151947
KEYWORDS   Euk.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1. (143451, 1-4). (220250)

```

AUTHORS Platzner M. and Varon R.
 TITLE Patent Submission
 JOURNAL Submitted (09.1997) Genome Analysis, Institute of Molecular
 Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
 COMMENT This sequence is part of a larger genomic contig. The start of this
 sequence is directed towards the centromere. The end of this
 sequence (32000..320250) overlaps with the start of the
 neighbouring Acc number AF049895 (1..250).
 FEATURES
 source location/qualifiers
 1..320250
 Organism "Homo sapiens"
 Zdb_xref "taxon:9606"
 Chromosome "8"
 Zmap "8q21.3"
 contig location/qualifiers
 1..137184
 Organism "Homo sapiens"
 Zdb_xref "taxon:9606"
 Zclone "P64M4"
 repeat_region complement(572..1704)
 Zrpt_family "MIR"
 repeat_region 205..2870
 Zrpt_family "MIR"
 repeat_region complement(946..1985)
 Zrpt_family "MIR"
 repeat_region 2222..2456
 Zrpt_family "L2"
 repeat_region 2538..2601
 Zrpt_family "L2"
 repeat_region complement(2695..2845)
 Zrpt_family "FAM"
 repeat_region 2913..3485
 Zrpt_family "L2"
 repeat_region 3559..3676
 Zrpt_family "L2"
 repeat_region 3726..3909
 Zrpt_family "MIR"
 repeat_region complement(4005..4247)
 Zrpt_family "MIR"
 repeat_region complement(5211..5555)
 Zrpt_family "L2"
 repeat_region 5667..6180
 Zrpt_family "L2"
 repeat_region 6800..6850
 Zrpt_family "(TA)n"
 repeat_region 8389..8709
 Zrpt_family "L2"
 repeat_region complement(9413..9761)
 Zrpt_family "ML1A1"
 repeat_region 10399..10473
 Zrpt_family "L2"
 repeat_region 11341..11507
 Zrpt_family "MIR"
 repeat_region complement(11525..11570)
 Zrpt_family "(TAA)n"
 repeat_region 11931..11985
 Zrpt_family "(CAAT)n"
 repeat_region complement(12051..12149)
 Zrpt_family "MER5C"
 repeat_region complement(12271..12700)
 Zrpt_family "MS1A"
 repeat_region 13544..14734
 Zrpt_family "MIR"
 repeat_region complement(15740..15847)
 Zrpt_family "MIR"
 repeat_region 16491..16598
 Zrpt_family "MIR"
 repeat_region 18183..18162
 Zrpt_family "MIR"
 repeat_region 18343..18606
 Zrpt_family "AluSx"
 repeat_region 19097..19075
 Zrpt_family "MER5L internal"
 repeat_region complement(19478..19560)

Zrpt_family "MIR"
 complement(23416..23710)
 Zrpt_family "AluSq"
 complement(24833..25025)
 Zrpt_family "MIR"
 complement(25333..25392)
 Zrpt_family "MER81"
 25969..26053
 Zrpt_family "L1M8"
 26367..26302
 Zrpt_family "L1M4"
 complement(26512..26684)
 Zrpt_family "AluJo"
 26685..26712
 Zrpt_family "(TA)n"
 complement(26713..26841)
 Zrpt_family "AluJo"
 complement(27456..27620)
 Zrpt_family "MER5A"
 complement(27724..28020)
 Zrpt_family "AluY"
 28347..28283
 Zrpt_family "AluY"
 28284..28311
 Zrpt_family "(TA)n"
 28312..28476
 Zrpt_family "AluY"
 29222..29518
 Zrpt_family "L1M2"
 29454..32583
 Zrpt_family "L1MA6"
 32636..33403
 Zrpt_family "L1MA6"
 33404..33705
 Zrpt_family "AluSe"
 33706..33960
 Zrpt_family "L1MA6"
 complement(34346..34614)
 Zrpt_family "MER8"
 34786..34873
 Zrpt_family "MIR"
 34839..35057
 Zrpt_family "L2"
 complement(35171..35382)
 Zrpt_family "AluYa5"
 35577..35746
 Zrpt_family "ML1B"
 complement(35747..36046)
 Zrpt_family "AluY"
 36047..36323
 Zrpt_family "ML1B"
 complement(36593..36761)
 Zrpt_family "MIR"
 complement(37858..37899)
 Zrpt_family "L2"
 38552..38895
 Zrpt_family "L1M2"
 38959..39811
 Zrpt_family "L1MA6"
 39885..39989
 Zrpt_family "L1M2"
 40242..44333
 Zrpt_family "L1M2"
 43134..43499
 Zrpt_family "ML1A1"
 43500..44075
 Zrpt_family "L1M2"
 44076..44391
 Zrpt_family "AluJo"
 44392..45758
 Zrpt_family "L1MA4A"
 complement(45759..46216)
 Zrpt_family "L1PA7"

Qy	1541	TCAGAAAGGATGAGGAGGATGATATAGGAGGAGATGATAGAGAGCAAAAGGAGGACATT	1600
Db	179446	TCAGAAAGGATGAGGAGGATGATATAGGAGGAGGAGATGATAGAGCAAAAGGAGGACATT	179505
Qy	1601	GTGAAGCAAAATGACAGAGGCGTGGCTTAAGCCAGTGGCTAGATGCCCTCTGTCCAGGGAC	1660
Db	179506	GTGAAGCAAAATGACAGAGGCGTGGCTTAAGCCAGTGGCTAGATGCCCTCTGTCCAGGGAC	179565
Qy	1661	TTGATCATGAGAGGAGGACTATGAACTTTGTAGTTCAGAGGCTACAGGGAGCTCAAAAGTC	1720
Db	179566	TTGATCATGAGAGGAGGACTATGAACTTTGTAGTTCAGAGGCTACAGGGAGCTCAAAAGTC	179625
Qy	1721	AGACAAATATCATAGACACATACGACATCCCAAGGACAGAAATTTGCCAAAGTTATAGTACAA	1780
Db	179626	AGACAAATATCATAGACACATACGACATCCCAAGGACAGAAATTTGCCAAAGTTATAGTACAA	179685
Qy	1781	AAATTCAGAAATACAAACCAAAATGGGCTTCCAGCGCTATCCCGGAGAAATACGTGGGTTCT	1840
Db	179686	AAATTCAGAAATACAAACCAAAATGGGCTTCCAGCGCTATCCCGGAGAAATACGTGGGTTCT	179745
Qy	1841	AGATGCCCATCTTAAATTTACTTCAGAAATAAAGGATGTAAGTGAGTGTCTTTCAGGAA	1900
Db	179746	AGATGCCCATCTTAAATTTACTTCAGAAATAAAGGATGTAAGTGAGTGTCTTTCAGGAA	179805

[illegible]

```

RESULT 1 (1)
AF194892
LOCUS       AF194892                1452 bp    mRNA       linear   PLN 92:AUG-2001
DEFINITION  Filobasidiella neoformans acetohydroxyacid reductoisomerase mRNA,
              complete cds.
ACCESSION   AF194892
VERSION     AF194892.1   GI:115077833
KEYWORDS    .
SOURCE      Filobasidiella neoformans.
            ORGANISM      Filobasidiella neoformans
                        Eukaryota; Fungi; Basidiomycota; Hymenomycetes;
                        Heterobasidiomycetes; Tremellomycetidae; Tremellales; Tremellaceae;
                        Filobasidiella.
REFERENCE   1  (bases 1 to 1452)
AUTHOR     McCusker, J.H.
TITLE      Identification of five amino acid biosynthesis genes from the human
            pathogen Cryptococcus neoformans
JOURNAL    Unpublished
REFERENCE   2  (bases 1 to 1452)
AUTHOR     McCusker, J.H.
TITLE      Direct Submission
JOURNAL    Submitted (26 JUN 2001) Microbiology, Duke University Medical

```

Matches	0;	Conservation	0;	Mismatches	0;	Indels	0;	Gaps	0;

Shovchenko, Y. I., Wetherby, K. D., Boekstyn, S. Sternberg, S. M.,

Search completed: May 13, 2002, 21:29:56
Job time: 8735 sec


```

RESULT 1
ID 4354
AC 04354 PRELIMINARY: PRI: 540 AA.
DE 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
TI SERINE-THREONINE KINASE RICK (RICK) (THREONINE-DEPENDENT) (SERINE-THREONINE
DE KINASE 2).
GN RICK (or RIP2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 98241596; PubMed 9575181;
RA Inohara N., del Peso L., Koseki T., Chen S., Ninez S.;
TI "RICK, a novel protein kinase containing a caspase recruitment domain,
RI interacts with CLARP and regulates CD95-mediated apoptosis.";
R1 J. Biol. Chem. 273:12299-12300(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Ozersky P., Holmes A., Broy M.;
PI Submitted (JAN 1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE 98307936; PubMed 9642260;
PA McCarthy J.V., Nijm J., Dixit V.M.;
TI "RIP2 is a novel NF-kappaB-activating and cell death-inducing
RI kinase.";
R1 J. Biol. Chem. 273:16998-16975(1998).
RN [4]
RP SEQUENCE FROM N.A.
RA Thome M., Holmann K., Burns K., Martinon F., Bodmer J.-L.,
PA Matsuda C., Tschopp J.;
TI "Identification of CARDIAK, a RIP-like kinase that associates with
RI caspase-1.";
R1 Curr. Biol. 8:1030(1998)

```

EN 1
 RP SEQUENCE FROM N.A.
 RA Pridmore M., Voth R.
 DE Submitted (DEC 1998) to the EMBL/Genbank/DBJ databases.
 GN 1
 OS SEQUENCE FROM N.A.
 OT TISSUE_KIN. AND MELANOMA;
 RA Strausberg R.
 DE Submitted (MAR 2001) to the EMBL/Genbank/DBJ databases.
 CC 1. SIM 1A111; PRO-NCN 1; RB. PRO-THE FAMILY OF PROTEIN KINASES;
 EMBL: A004003; AA034900.1; 1;
 EMBL: A004003; AA024601.1; 1;
 EMBL: A004003; AA027722.1; 1;
 EMBL: A004003; AA025666.1; 1;
 EMBL: A004003; AA004034.1; 1;
 EMBL: A004003; AA004553.1; 1;
 DE InterPro: IPR001315; PAKD;
 DE InterPro: IPR00719; Euk_pkinase;
 DE InterPro: IPR02499; Ser_thr_pkinase;
 DE InterPro: IPR04454; STY_pkinase;
 DE Pfam: PF00069; PAKD; 1;
 DE Pfam: PF00069; pkinase; 1;
 DE SMART: SM0114; PAKD; 1;
 DE SMART: SM012; STYK; 1;
 DE PROSITE: PS00099; PAKD; 1;
 DE PROSITE: PS0011; PROTEIN_KINASE_DOM; 1;
 DE PROSITE: PS00100; PROTEIN_KINASE_ST; 1;
 FW ATP-binding; kinase; Serine/threonine; protein kinase; transferase
 520NN 54 AA: 6104 MW: 75A022605792 chr64;

DE 01 JAN 1998 (TrEMBLrel. 05, last sequence update)
 DE 01 DEC 2001 (TrEMBLrel. 19, last annotation update)
 DE PROTEIN TYROSINE KINASE (EC 2.7.1.112) (FRAGMENT).
 GN 8;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Telodoridae; Caenorhabditis.
 OX NCBI_TaxID:6239;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Gates A.C., Wilks A.F.;
 RT "How many PKs to organise a worm?";
 RL The Worm Breeders Gazette 14:87-87(1995).
 RN 12
 RP SEQUENCE FROM N.A.
 RA Gates A.C., Wollberg P., Achen M.G., Wilks A.F.;
 RT "Sampling the Genomic Pool of Protein Tyrosine Kinase Genes using the
 RT Polymerase Chain Reaction with Genomic DNA.";
 RL Submitted (NOV 1997) to the EMBL/Genbank/DBJ databases
 DE EMBL: A1002899; CAA05734.1; 1;
 DE InterPro: IPR000719; Euk_pkinase;
 DE Pfam: PF00069; pkinase; 1;
 DE PROSITE: PS00011; PROTEIN_KINASE_DOM; 1;
 KW ATP-binding; Kinase; Transferase;
 FI NON_TER 1 1
 FI NON_TER 54 54
 SQ SEQUENCE 54 AA: 6030 MW: 6P36PR2306580; CR64;

Query Match: 1.00; Score 10; DB 5; Length 54;
 Best Local Similarity: 100.0%; Pred. No. 0.01;
 Matches: 10; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

Query Match: 85.4%; Score 461; DB 4; Length 540;
 Best Local Similarity: 100.0%; Pred. No. 0;
 Matches: 41; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

02 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 141 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 12 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 13 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 14 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 15 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 16 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 17 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 18 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 19 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 20 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 21 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 22 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 23 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 24 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 25 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 26 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 27 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 28 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 29 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 30 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 31 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 32 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 33 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 34 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 35 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 36 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 37 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 38 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 39 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 40 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 41 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 42 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 43 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 44 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 45 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 46 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 47 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 48 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 49 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 50 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 51 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 52 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 53 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 54 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 55 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 56 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 57 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 58 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 59 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 60 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 61 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 62 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 63 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 64 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 65 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 66 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 67 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 68 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 69 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 70 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 71 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 72 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 73 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 74 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 75 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 76 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 77 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 78 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 79 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 80 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 81 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 82 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 83 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 84 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 85 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 86 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 87 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 88 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 89 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 90 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 91 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 92 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 93 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 94 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 95 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 96 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 97 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 98 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 99 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199
 100 140 PGLDNEEFGLIVTYEYMGNSLNELIHKKTEYFVAVWFLERLRLHETALGVNYLHMT 199

EN 1
 RP SEQUENCE FROM N.A.
 RA Pridmore M., Voth R.
 DE Submitted (DEC 1998) to the EMBL/Genbank/DBJ databases.
 GN 1
 OS SEQUENCE FROM N.A.
 OT TISSUE_KIN. AND MELANOMA;
 RA Strausberg R.
 DE Submitted (MAR 2001) to the EMBL/Genbank/DBJ databases.
 CC 1. SIM 1A111; PRO-NCN 1; RB. PRO-THE FAMILY OF PROTEIN KINASES;
 EMBL: A004003; AA034900.1; 1;
 EMBL: A004003; AA024601.1; 1;
 EMBL: A004003; AA027722.1; 1;
 EMBL: A004003; AA025666.1; 1;
 EMBL: A004003; AA004034.1; 1;
 EMBL: A004003; AA004553.1; 1;
 DE InterPro: IPR001315; PAKD;
 DE InterPro: IPR00719; Euk_pkinase;
 DE InterPro: IPR02499; Ser_thr_pkinase;
 DE InterPro: IPR04454; STY_pkinase;
 DE Pfam: PF00069; PAKD; 1;
 DE Pfam: PF00069; pkinase; 1;
 DE SMART: SM0114; PAKD; 1;
 DE SMART: SM012; STYK; 1;
 DE PROSITE: PS00099; PAKD; 1;
 DE PROSITE: PS0011; PROTEIN_KINASE_DOM; 1;
 DE PROSITE: PS00100; PROTEIN_KINASE_ST; 1;
 FW ATP-binding; kinase; Serine/threonine; protein kinase; transferase
 520NN 54 AA: 6104 MW: 75A022605792 chr64;

02 140 VKIADPGLSK 18
 141 VKIADPGLSK 18

RESULT 1
 Q9D685
 ID Q9D685 PRELIMINARY; PRI; 291 AA.
 AC Q9D685;
 DT 01 JUN 2001 (TrEMBLrel. 17, Created)
 DT 01 JUN 2001 (TrEMBLrel. 17, last sequence update)
 DT 01 OCT 2001 (TrEMBLrel. 18, last annotation update)
 DE 1632401P23R1K; PROTEIN;
 GN 4632401P23R1K;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurimorphi; Muridae; Murinae; Mus;
 OX NCBI_TaxID:10090;
 RN 11
 RP SEQUENCE FROM N.A.
 RA STRAIN C57BL/6J; TISSUE SKIN;
 RX MEDLINE 21085660; PubMed 11217851;
 RA Nakai H., Shimada A., Shibata K., Yoshida H., Itoh H., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kamei H., Adachi J., Fukuda S.,
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Sando S., Yamakura T.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito K.,
 RA Tadota K., Matsuda H.A., Ashburner M., Ratliff J., Charoyant T.,
 RA Fleischmann W., Gausterland T., Gissi C., King R., Kochiwa H.,
 RA Fiedler G., Lewis J., Matsuo Y., Nakai H., Saito T., Saito K., Saito Y.,
 RA Schiraldi L.M., Staubli P., Suzuki K., Tomita M., Warner H., Washio T.,
 RA Akai H., Akai T., Furuta M., Zeng H., Baldridge J., Barish G.,
 RA Blake J., Bolte H., Bojunga N., Carninci P., de Bonis M.F.,
 RA Branstetter M.J., Bull J., Fletcher S., Fujita H., Gariboldi M.,
 RA Gasteiger S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli L., Mombaerts P.,
 RA Nordone P., King R., Rinaldi M., Rodriguez L., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Soya T., Shibata Y., Storch K.,
 RA Suzuki H., Toyooka K., Wang K.B., Wells C., Whitaker C., Wilmut L.,
 RA Wynshaw-Boris A., Yoshida K., Yasuoka Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 CC -1: SIMILARITY: BLASTX TO THE SWISS-PROT FAMILY OF PROTEIN KINASES.
 DR EMBL: AK014746; BAB29424.1; -;
 DR MGI: MGI:1921589; 4632401F23kik.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00220; STKc; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 291 AA; 42902 MW; DF2B7523A467021F CRC64;

Query Match 1.94; Score 10; DB 11; Length 291;
 Best Local Similarity 100.0%; Pred. No. 0.077;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 BLKQNI111D 155
 DB 128 BLKQNI111D 147

RESULT 4
 Q21292
 ID Q21292 PRELIMINARY; PRT; 517 AA.
 AC Q21292;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE K07F5.4 PROTEIN.
 GN K07F5.4
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Thrematodea; Eladitida; Eladitidae.
 CC Eladitidae; Poloderinae; Caenorhabditis.
 OX NCBI_TaxID 6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hombry C;
 PI Submitted (MAR 1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 9506613; PubMed 9501916.
 RA Jones;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology.";
 RL Science 282:2162-2168(1998).
 DR EMBL: Z70284; CAA94277.1; -;
 DR HSSP: P11342; 1FGK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00001; SH2; 1.
 KW ATP-binding; Transferase; Tyrosine-protein kinase.
 SQ SEQUENCE 517 AA; 58618 MW; A82974BD4CABCF7 CRC64;

Query Match 1.9%; Score 10; DB 5; Length 517;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VETAFNLSK 169
 DB 266 VETAFNLSK 169

RESULT 5
 O18645
 ID O18645 PRELIMINARY; PRT; 582 AA.
 AC O18645; Q9U4G0;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE DETA PROTEIN (DETA PROTEIN).
 GN DETA OR EG14208.2 OR CG3051.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Ephydrophila; Drosophilidae; Drosophila.
 OX NCBI_TaxID 7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN PEGASLEY;
 RX MEDLINE 20196096; PubMed 10741142;
 RA Adams M.D., Schirmer S.R., Heit K.A., Evans C.A., Isayane J.D.,
 RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Watanabe J.R., Yandell M.B., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.B.,
 RA Wan K.H., Doyle P., Baxter E.G., Heit K., Nelson C.K., Miklos G.L.G.,
 RA Abril J.F., Ashayari A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballow R.M., Basu A., Boxendale L., Boyrakatoglou L., Bozley E.M.,
 RA Brown K.Y., Brown P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.E., Botch G., Brockton P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Chesley P., Dalke C., Davenport T.R., Davies P.,
 RA de Babblos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson E., Doup E.E., Downes M., Duran-Rocha S., Dunkov R.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Easler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hartzel H., Hickey J., Holman J., Hornig M., Hornig J., Hunk J.,
 RA Huxlin D., Houston K.A., Howland T.J., Wei M.-H., Iboqwan C.,
 RA Izatt M., Kalish J., Karpen J.H., Ke Y., Kien J.A., Kien J.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lee Y., Levitsky A.A., Li J., Li Z., Liang Y., Liu X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markakis C., Melnick M.V., Melnik S., Miller S., Mitchell A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,
 RA Nelson D.E., Nelson K.A., Nison K., Nusskera D.K., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Poulard D., Puri V., Reese M.G.,
 RA Reinert K., Rombauten K., Saunders R.D.C., Scheeler P., Shen H.,
 RA Shao R.C., Siden-Kiamos I., Simpson M., Skapaki M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Stroud R., Sun E.,
 RA Svirkas K., Tector C., Turner K., Venter R., Wang A.H., Wang X.,
 RA Wang Z. Y., Warsman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodard L., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Ye K., Ye L., Ye L., Ye L., Ye L., Ye L., Ye L., Ye L.,
 RA Zhang Y.H., Zhang P.H., Zhang W., Zhang X., Zhang X., Zhang X.,
 RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.L.
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN PEGASLEY;
 RA Taborda E.N., Hickey D.A.;
 RL Submitted (JUN 1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN PEGASLEY;
 RA Taborda E.N., Hickey D.A.;
 RL Submitted (JUN 1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN PEGASLEY;
 RA Papadogiannakis S., Spanos I., Cox S., Siden Kiamos I., Louis C.;
 RL Submitted (JUN 1998) to the EMBL/GenBank/DBJ databases.

RN [1]
 RF SEQUENCE FROM N.A.
 RC STRAIN BERKELEY
 RA Bobbit L.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
 Adayar A., Arriaga L.L., Baxter E., Blazej R., Boteschell C.,
 Champe M., Chavez C., Hew M., Hylleberg M., Fartan D., Frise R.,
 Galle R., George K.A., Harris N.L., Hoskins P.A., Evans-Bohm M.,
 Hsu K.A., Hummel S.R., Kim E., Li P., Moshrefi M., Paclch J.M.,
 Park S., Sepurwa A., Sethi B., Ruiz E., Svirskas R.E., Weinberg L.,
 Wilson J.
 RI "Full length *Drosophila melanogaster* cDNA sequence."
 RL Submitted (AUG 1999) to the EMBL/GenBank/DBJ databases.
 SC 1 SIMILARITY: BELONG TO THE SER/THR FAMILY OF PROTEIN KINASES
 DR EMBL: AF004120; AAF45614.1
 DR EMBL: AF020110; AAF71498.1
 DR EMBL: AF020109; AAF71497.1
 DR EMBL: AF020990; CAA19654.1
 DR EMBL: AF181479; AAD95495.1
 DR HSSP: P03450; IAP9.
 DR FlyBase: FBan002169; GNFA.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002200; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase_1.
 DR SMART: SM00221; STYKc_1.
 DR PROSITE: PS00407; PROTEIN_KINASE_ATP_1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST_1.
 KW ATP binding; Serine/threonine protein kinase; Transferase.
 SQ SEQUENCE 678 AA; 64599 MW; 8D6B84E6B627B67C CR664;

Query Match 1.99%; Score 10; DB 5; Length 582;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 116 WEVLSRKQPF 198

DI 116 WEVLSRKQPF 198

DB 116 WEVLSRKQPF 198

RESULT 6

ID Q9VXQ2 PRELIMINARY; PRI: 678 AA;
 AC Q9VXQ2
 DI 01 MAY 2000 (1REMBLrel. 13; Created)
 DI 01 MAY 2000 (1REMBLrel. 13; Last sequence update)
 DI 01 DEC 2001 (1REMBLrel. 19; Last annotation update)
 DE INITIATIVE TAKI PROTEIN (Q9VXQ2 PROTEIN) (U422741).
 IN TAKI (c. 678aa) OR C618492.
 OR *Drosophila melanogaster* (Fruit fly).
 OR Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OR Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OR Ephydrales; Drosophilidae; *Drosophila*.
 OR NCBI TaxId 7227;
 RN [1]
 RF SEQUENCE FROM N.A.
 RC STRAIN BERKELEY
 RX MEDLINE 2019506; PubMed 10731142;
 RA Angus R.E., Celisner S.E., Holt K.A., Evans C.A., Gocayne J.D.,
 Ananif D.P., Bohrer S.E., Li P.W., Hoskins P.A., Galle R.E.,
 George K.A., Lewis S.E., Richards R., Ashburner M., Henderson S.N.,
 Patton J.L., Warrman J.R., Vardell M.J., Zhang G., Chen L.X.,
 Brandon R.C., Rogers Y.B.C., Blazej R., Champe M., Pfeiffer B.D.,
 Wan K., Taylor C., Baxter E., Holt G., Nelson C.E., Miklos G.L.G.,
 Abell C., Aibayant A., An H., Andrews Pankovich C., Baldwin D.,
 Bellow S.M., Busch A., Baxendale J., Bayraktarova L., Beasley K.M.,
 Beson R.Y., Boman P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Borokov A., Botchan M.R., Boeck J., Brokstein P., Brothier P.,
 Burtis K.L., Busan D.A., Butler H., Cadieu E., Center A., Chandra L.,
 Cherry M., Dawley S., Dalke C., Davenport L.K., Davies P.,
 de Tilly R., Delcher A., Deng Z., Mays A.D., Hew L., Dietz S.M.,
 DeLeon S., Deep L.E., Downen M., Duran Rocha S., Dunkov R.C., Dunn P.,
 Durbin K.L., Egan J., Farnham J., Ferraz C., Ferrier S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garq N.S., Gelbart W.M., Glasser K.,
 Glodok A., Goni F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hustin D., Houston K.A., Howland T.J., Wei M.H., Hsuwam C.,
 Jalali M., Kalish R., Karpen G.H., Ko Z., Kohnsen J.A., Kotchum K.A.,
 Kimmel P.K., Kohnen C., Kravitz S., Kulp D., Lal Z.,
 Laske R., Lei Y., Lowitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy B., Murphy B., Murphy L., Muny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon E., Nusskern D.K., Paclch J.M.,
 Palamolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,
 Roberts E., Roelstra F., Sanders R.D., Scheller P., Shen H.,
 Shue R.C., Siden Kianes L., Simpson M., Skopski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong K., Sun E.,
 Svirskas R., Teeter C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z., Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zhang X.H., Zhang P.N., Zhang W., Zhou X., Zhu X., Zhu X., Smith H.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter L.C.
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RF SEQUENCE FROM N.A.
 RA Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
 O'Connor M.B., Shibuya H., Hono N.
 SC Submitted (1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RF SEQUENCE FROM N.A.
 RC STRAIN Y, CN HW SH;
 RA Stapleton M., Brokstein P., Bond L., Adayar A., Arriaga J.,
 Champe M., Chavez C., Dorsett V., Fartan D., Frise R., George K.,
 Gonzalez M., Guarin R., Li P., Liao G., Miranda A., Munhall C.J.,
 Nanuo J., Paclch J., Paragas V., Park S., Phouanavong S., Wan K.,
 Yu C., Lewis S.E., Rubin G.M., Celisner S.
 RI Submitted (AUG 2001) to the EMBL/GenBank/DBJ databases.
 SC 1 SIMILARITY: BELONG TO THE SER/THR FAMILY OF PROTEIN KINASES
 DR EMBL: AF004571; AAF50895.1
 DR EMBL: AF199466; AAF06815.1
 DR EMBL: AY051953; AAK94377.1
 DR HSSP: P08631; IAD5.
 DR FlyBase: FBan002643; Taki.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001990; Granin.
 DR InterPro: IPR002200; Ser_thr_kinase.
 DR InterPro: IPR004040; STY_pkinase.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase_1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00221; STYKc_1.
 DR PROSITE: PS00422; GRANINS_1; UNKNOWN_1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP_1.
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM_1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST_1.
 KW ATP binding; Hypothetical protein; Kinase;
 KW Serine/threonine protein kinase; Transferase.
 SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDB8C1045 CR664;

Query Match 1.99%; Score 10; DB 5; Length 678;
 Best Local Similarity 100.0%; Pred. No. 0.16;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 WEVLSRKQPF 229

DI 220 WEVLSRKQPF 229

DB 220 WEVLSRKQPF 213

RESULT 7

ID Q9VXQ2 PRELIMINARY; PRI: 678 AA;
 AC Q9VXQ2

DI 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE NIMA-RELATED KINASE B.
 GN NEK8.
 OS Brachydanio rerio (Zebrafish) (Sebida Jario).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi.
 CC Cypriniformes; Cyprinidae; Danio.
 CX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu S., Lu W., Obara-Ishihara T., Drummond L., Beier D.K.;
 RT "A defect in a novel Nek-family kinase causes cystic disease in the
 RI mouse and in zebrafish."
 PI Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases
 DR EMBL: AF407580; AAL09676.1;
 KW Kinase.
 SQ SEQUENCE 697 AA; 76541 MW; 875A51D1E3831AA7 CRC64;

Query Match 1.9%; Score 10; DB 13; Length 697;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 DLKIQNILLD 155
 |||
 DB 128 DLKIQNILLD 137

RESULT 8
 Q91ZR4
 ID Q91ZR4 PRELIMINARY; PRT; 698 AA.
 AC Q91ZR4;
 DI 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE NIMA-RELATED KINASE B.
 GN NEK8.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Liu S., Lu W., Obara-Ishihara T., Drummond L., Beier D.K.;
 RT "A defect in a novel Nek family kinase causes cystic disease in the
 RI mouse and in zebrafish."
 PI Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF407579; AAL09675.1;
 KW Kinase.
 SQ SEQUENCE 698 AA; 75264 MW; 686B29A8C7180E94 CRC64;

Query Match 1.9%; Score 10; DB 11; Length 698;
 Best Local Similarity 100.0%; Pred. No. 0.17;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 DLKIQNILLD 155
 |||
 DB 128 DLKIQNILLD 137

RESULT 9
 Q9GJS5
 ID Q9GJS5 PRELIMINARY; PRT; 22 AA.
 AC Q9GJS5;
 DI 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DI 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE CA2/CALMODULIN INDEPENDENT KINASE (FRAGMENT).
 OS Lagenorhynchus obliquidens (Pacific white-sided dolphin).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 CC Lagenorhynchus.
 CX NCBI_TaxID=90247;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN: CCB_141R, AND CCB_141S;
 RA Haro M.P., Cipriano P., Palumbi S.R.;
 RT "Slow Evolution of Genetic Monophyly in Dolphins: Implications for
 RT Speciation, Systematics and Conservation."
 PI Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF140810; AAF98663.1;
 DR EMBL: AF140809; AAF98662.1;
 KW Kinase.
 FI NON-ITER 1 1
 FI NON-ITER 2 2
 SQ SEQUENCE 22 AA; 2346 MW; A2A16ACBDAF3DD0 CRC64;

Query Match 1.7%; Score 9; DB 6; Length 22;
 Best Local Similarity 100.0%; Pred. No. 0.09;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KIADEGLSK 169
 |||
 DB 8 KIADEGLSK 16

RESULT 10
 Q9GJW0
 ID Q9GJW0 PRELIMINARY; PRT; 33 AA.
 AC Q9GJW0;
 DI 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DI 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CA2/CALMODULIN INDEPENDENT KINASE (FRAGMENT).
 OS Lagenorhynchus obscurus (dusky dolphin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 CC Lagenorhynchus.
 CX NCBI_TaxID=27611;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA Haro M.P., Cipriano P., Palumbi S.R.;
 RT "Slow Evolution of Genetic Monophyly in Dolphins: Implications for
 RT Speciation, Systematics and Conservation."
 PI Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF140820; AAF98673.1;
 DR EMBL: AF140813; AAF98666.1;
 DR EMBL: AF140814; AAF98667.1;
 DR EMBL: AF140817; AAF98670.1;
 DR EMBL: AF140816; AAF98671.1;
 DR EMBL: AF140819; AAF98672.1;
 DR InterPro: IPR000719; Euk_pkinase
 DR Pfam: PF00069; pkinase; 1.
 DR PROSITE: PS00811; PROTEIN_KINASE_DOM; 1.
 KW ATP binding, Kinase; Transferase.
 FI NON-ITER 1 1
 FI NON-ITER 33 33
 SQ SEQUENCE 33 AA; 3604 MW; 0664A909F350D001 CRC64;

Query Match 1.7%; Score 9; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KIADEGLSK 169
 |||
 DB 19 KIADEGLSK 27

RESULT 11

Q9GJV9 PRELIMINARY: PRI: 33 AA.
 AT Q9GJV9;
 DI 01-MAR-2001 (TrEMBLrel. 16, Created)
 DI 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DI 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CA2/CALMODULIN DEPENDENT KINASE (FRAGMENT).
 OS Lagenorhynchus obliquidens (Pacific white-sided dolphin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 OC Lagenorhynchus.
 OX NCBI TaxID: 96047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN LAC_1230_B, AND LAC_1239;
 RA Hare M.P., Cipriano F., Palumbi S.R.;
 RL "Slow Evolution of Genetic Monophyly in Dolphins: Implications for
 Speciation, Systematics and Conservation.";
 RT Submitted (Apr 1999) to the EMBL/GenBank/DBS databases.
 LR EMBL: AF140804; AAF98659.1;
 LR EMBL: AF140804; AAF98659.1;
 LR EMBL: AF140804; AAF98659.1;
 LR EMBL: AF140804; AAF98659.1;
 DR InterPro: IPR000719; Euk_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DE PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP binding; Kinase; Transferase.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 6604 MW; 0604A999F5C6E01 CRC64;
 Query Match 1.7%; Score 9; DB 6; Length 33;
 Best Local Similarity 100.0%; Prod. No. 0.13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 01-KINDEGLSK 169
 DI 01-KINDEGLSK 27

RESULT 14
 Q9GMB4 PRELIMINARY: PRI: 33 AA.
 AT Q9GMB4;
 DI 01-MAR-2001 (TrEMBLrel. 16, Created)
 DI 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DI 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CA2/CALMODULIN DEPENDENT KINASE (FRAGMENT).
 OS Lagenorhynchus obliquidens (Pacific white-sided dolphin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 OC Lagenorhynchus.
 OX NCBI TaxID: 96047;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN LAC_1230_B, AND LAC_1239;
 RA Hare M.P., Cipriano F., Palumbi S.R.;
 RL "Slow Evolution of Genetic Monophyly in Dolphins: Implications for
 Speciation, Systematics and Conservation.";
 RT Submitted (Apr 1999) to the EMBL/GenBank/DBS databases.
 LR EMBL: AF140804; AAF98660.1;
 LR EMBL: AF140804; AAF98660.1;
 LR InterPro: IPR000719; Euk_pkinase.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP binding; Kinase; Transferase.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 6601 MW; 0611FD1CA75C0E01 CRC64;
 Query Match 1.7%; Score 9; DB 6; Length 33;
 Best Local Similarity 100.0%; Prod. No. 0.13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 161 KIADEGLSK 169
 11111111
 LB 19 KIADEGLSK 27

RESULT 14
 Q9GJV7 PRELIMINARY: PRI: 33 AA.
 AT Q9GJV7;
 DI 01-MAR-2001 (TrEMBLrel. 16, Created)
 DI 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DI 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CA2/CALMODULIN DEPENDENT KINASE (FRAGMENT).
 OS Lagenorhynchus obscurus (dusky dolphin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 OC Lagenorhynchus.
 OX NCBI TaxID: 27611;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN LOS_236_B, AND LOS_236;
 RA Hare M.P., Cipriano F., Palumbi S.R.;
 RL "Slow Evolution of Genetic Monophyly in Dolphins: Implications for
 Speciation, Systematics and Conservation.";
 RT Submitted (Apr 1999) to the EMBL/GenBank/DBS databases.
 LR EMBL: AF140816; AAF98669.1;
 LR EMBL: AF140816; AAF98669.1;
 DR InterPro: IPR000719; Euk_pkinase.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP binding; Kinase; Transferase.
 FT NON_TER 1 1
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 6598 MW; 0611FD1CA75C0E01 CRC64;

Query Match 1.7%; Score 9; DB 6; Length 33;
 Best Local Similarity 100.0%; Prod. No. 0.13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KIADEGLSK 169
 11111111
 LB 19 KIADEGLSK 27

RESULT 14
 Q9GMB4 PRELIMINARY: PRI: 33 AA.
 AT Q9GMB4;
 DI 01-MAR-2001 (TrEMBLrel. 16, Created)
 DI 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DI 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE CA2/CALMODULIN DEPENDENT KINASE (FRAGMENT).
 OS Lagenorhynchus acutus (Atlantic white-sided dolphin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti; Delphinidae;
 OC Lagenorhynchus.
 OX NCBI TaxID: 90246;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN LAC_9400;
 RX MEDLINE: 20051255; PubMed 10583838;
 RA Hare M.P., Palumbi S.R.;
 RL "The accuracy of heterozygous base calling from diploid sequence and
 resolution of haplotypes using allele-specific sequencing.";
 RT Mol. Ecol. 8:1750-1752(1999).
 LR EMBL: AF140821; AAF98674.1;
 DR InterPro: IPR000719; Euk_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP binding; Kinase; Transferase.
 FT NON_TER 1 1

FI NON_TER 33 33
 SQ SEQUENCE 33 AA: 3604 MW: 0604A909F35D0E01 CRC64;

Query Match 1.7%; Score 9; L3 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 0.14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 KIADEMLSK 169
 11111111
 DB 19 KIADEMLSK 27

RESULT 15
 Q38987
 ID Q38987 PRELIMINARY: PRI: 55 AA.
 AC Q38987;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE PROTEIN KINASE CATALYTIC DOMAIN (FRAGMENT).
 GN AK21.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_taxID 3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN EIL 0; TISSUE LEAF;
 KX MEDLINE 96124234; PubMed 8534852;
 KA Thuemmler F., Kirchner M., Teuber R., Dietrich P.;
 FI "Differential accumulation of the transcripts of 22 novel protein
 KI kinase genes in Arabidopsis thaliana.";
 KL Plant Mol. Biol. 29:551-565(1995).
 DR EMBL: X86966; CAA60529.1; -;
 DR InterPro: IPR000719; Euk_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW ATP-binding; Kinase; Transferase.
 FI NON_TER 1 1
 SQ SEQUENCE 55 AA: 5910 MW: 0E55D6DBD25A9FB CRC64;

Query Match 1.7%; Score 9; L3 10; Length 55;
 Best Local Similarity 100.0%; Pred. No. 0.2;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADFGLS 168
 11111111
 DB 13 VKIADFGLS 21

Search completed: May 14, 2002, 21:37:13
 Job time: 368 sec



GenCore version 1.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

9M protein protein search, using sw model

Run on: May 13, 2002, 21:30:00 ; Search time 33.25 seconds
(without alignments)
626.829 Million cell updates/sec

Title: US-09-445-223-1
Perfect score: 540
Sequence: 1 MNGEATGALPTPIYHKLAD...PEILVVSRSPLNLLQNKSM 540

Scoring table: OLIGO
Gapop 60.0 ; Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size: 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database: SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% query Match	Length	DB ID	Description
1	9	1.7	448	1	A2AR_HUMAN
2	9	1.7	469	1	KCC4_MOUSE
3	9	1.7	473	1	KCC4_HUMAN
4	9	1.7	474	1	KCC4_RAT
5	9	1.7	512	1	K110_ARATH
6	9	1.7	576	1	SNF1_SCHER
7	9	1.7	611	1	SNF1_CANGA
8	9	1.7	619	1	SNF1_CANTH
9	9	1.7	620	1	SNF1_CANAL
10	9	1.7	633	1	SNF1_YEAST
11	9	1.7	1100	1	JAK3_HA
12	9	1.7	1124	1	JAK3_HUMAN
13	9	1.7	1200	1	JAK3_MOUSE
14	8	1.5	74	1	CSME_CHLTI
15	8	1.5	99	1	P1XX_RHIZM
16	8	1.5	132	1	AAK1_PIG
17	8	1.5	174	1	AAK1_HUMAN
18	8	1.5	287	1	PKI1_MOUSE
19	8	1.5	295	1	KMLC_DICB
20	8	1.5	311	1	CC23_TRYB
21	8	1.5	342	1	BMAP_HUMAN
22	8	1.5	513	1	DUN1_YEAS
23	8	1.5	549	1	AAK1_RAT
24	8	1.5	552	1	AAK2_HUMAN
25	8	1.5	552	1	AAK2_RAT
26	8	1.5	622	1	MAK_MOUSE
27	8	1.5	622	1	MAK_RAT
28	8	1.5	622	1	YNA3_CAEEL
29	8	1.5	624	1	MAK_HUMAN
30	8	1.5	639	1	KPCI1_MOUSE
31	8	1.5	661	1	Y537_HUMAN
32	8	1.5	842	1	PKI2_RAT
33	8	1.5	923	1	10R_DROME

44	8	1.5	942	1	PKI1_HUMAN	Q16512 homo sapien
45	8	1.5	946	1	PKI1_RAT	Q63133 rattus norv
46	8	1.5	980	1	CLV1_ARATH	Q55498 arabidopsis
47	8	1.5	984	1	PKI2_HUMAN	Q16513 homo sapien
48	8	1.5	999	1	RLK5_ARATH	P47735 arabidopsis
49	8	1.5	1122	1	TIE2_MOUSE	Q02858 mus musculu
40	8	1.5	1124	1	TIE2_HUMAN	Q02763 homo sapien
41	8	1.5	1124	1	TIE2_BOVIN	Q06807 bos taurus
42	8	1.5	1144	1	TIE1_MOUSE	Q06806 mus musculu
43	8	1.5	1144	1	TIE1_BOVIN	Q06805 bos taurus
44	8	1.5	1138	1	TIE1_HUMAN	P35590 homo sapien
45	8	1.5	1196	1	ABL1_CAEEL	P03949 caenorhabdi

ALIGNMENTS

RESULT 1					
AZAB_CAVPO					
ID	A2AR_CAVPO	STANDARD:	PKI:	448 AA.	
AC	Q60475:				
DT	01 NOV 1997 (Rel. 35, Created)				
DI	01 NOV 1997 (Rel. 35, Last sequence update)				
DI	16 OCT 2001 (Rel. 40, Last annotation update)				
DE	Alpha 2B adrenergic receptor (Alpha-2B adrenoceptor).				
GN	ADRA2B.				
OS	Cavia porcellus (Guinea pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Hysticomathi; Caviidae; Cavia.				
OX	NCBI:taxid:10141;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN: HARTLEY;				
RX	EMBL: U615257; PubMed 8573196;				
RA	Svensson S.P., Bailey L.J., Porter A.C., Ricman J.G., Regan J.W.;				
PT	"Heterologous expression of the cloned guinea pig alpha 2A, alpha 2B,				
RT	and alpha 2C adrenoceptor subtypes. Radioligand binding and				
RT	functional coupling to a CAMP-responsive reporter gene.";				
RI	Biochem. Pharmacol. 51:291-300(1995).				
CC	FUNCTION: ALPHA 2 ADRENERGIC RECEPTOR MEDIATE THE CATECHOLAMINE-				
CC	INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G				
CC	PROTEINS.				
CC	FUNCTION: LOCALIZATION: Integral membrane protein.				
CC	SIMILARITY: BELONGS TO FAMILY 1 OF G PROTEIN COUPLED RECEPTORS.				
CC				
CC	This SWISS PRO entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	enterprises requires a license agreement. For more information, please				
CC	or send an email to license@ebi.ac.uk.				
CC				
DR	EMBL: U615257; AAA67075.1;				
DR	InterPro: IPR003276; GPCR_Rhodopsn.				
DR	Pfam: PF00001; 7tm1; 1.				
DR	PROSITE: PROSITE: GPCR_Rhodopsn.				
DR	PROSITE: PS00247; G_PROTEIN_RECEP_F1_1; 1.				
DR	PROSITE: PS00247; G_PROTEIN_RECEP_F1_2; 1.				
KW	G protein coupled receptor; Transmembrane; Multitene family;				
KW	Phosphorylation; Lipoprotein; Palmitate.				
FT	DOMAIN 1 12				EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 13 38				1 (POTENTIAL).
FT	DOMAIN 39 49				CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 50 75				2 (POTENTIAL).
FT	DOMAIN 76 85				EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 86 108				1 (POTENTIAL).
FT	DOMAIN 109 130				CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 131 153				1 (POTENTIAL).
FT	DOMAIN 154 168				EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 169 192				5 (POTENTIAL).
FT	DOMAIN 193 220				CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 221 291				6 (POTENTIAL).

E1 DOMAIN 395 403 EXTRACELLULAR (POTENTIAL).
 E1 TRANS MEM 181 187 (POTENTIAL).
 E1 DOMAIN 428 448 CYTOPLASMIC (POTENTIAL).
 E1 DISUB 165 166 BY SIMILARITY.
 E1 LPE1 446 446 PALMITATE (POTENTIAL).
 E1 DOMAIN 294 309 ASP/THR RICH (ACTIVIC).
 E1 SITE 125 125 IMPLICATED IN CATIONIC AMINO ACID BINDING (BY SIMILARITY).
 E1 SITE 125 125 IMPLICATED IN CATIONIC AMINO ACID BINDING (BY SIMILARITY).
 E1 SITE 125 125 IMPLICATED IN CATIONIC AMINO ACID BINDING (BY SIMILARITY).
 SQ SEQUENCE 118 AA: 49597 MW: 8364187571344777 (K664)

Query Match 1.7% Score 9: DB 1: Length 448;
 Best Local Similarity 100.0% Pred. No. 0.45;
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 990 DBE1A92 998

DB 990 DBE1A92 47

RESULT 3

K0C4 HUMAN

ID K0C4 HUMAN STANDARD; PRI: 473 AA;
 AC Q16566;
 DT 01-NOV-1997 (Rel. 35, Created);
 DI 16-NOV-2001 (Rel. 40, Last sequence update);
 DE Calcium/calmodulin dependent protein kinase type IV catalytic chain
 (C47311.1.1.0) (CAM kinase-GK) (CaMK IV) [Contains: Caldesmonin].
 GN CAMK4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC BRAIN, BALB/c TISSUE Brain;
 RX MEDLINE 9433096; PubMed 1993667;
 RA Jones C.A., Reed L., Wilson Shaw D., Bahn W.E., Sikela J.M.;
 RT "cDNA sequence and differential expression of the mouse
 R1 122Z/calmodulin dependent protein kinase IV gene";
 R1 EPRS 1993; 1993;105:109(1993).
 RN [2]
 RP SEQUENCE OF 473 AA FROM N.A.
 RX MEDLINE 9423027; PubMed 250664.
 RA Sikela J.M., Luk M.L., Kao P.L., Harte J.A., Wei Q., Bahn W.E.;
 RT "Clonal localization of the human gene for brain
 R1 122Z/calmodulin dependent protein kinase type IV";
 R1 Science 264:227-227(1999).
 RN [3]
 RP SEQUENCE OF 473 AA FROM N.A.
 RC TISSUE Brain;
 RX MEDLINE 9429426; PubMed 944077;
 RA Sikela J.M., Bahn W.E.;
 RT "Screening an expression library with a ligand probe: isolation and
 R1 sequence of a cDNA corresponding to a brain calmodulin binding
 R1 protein";
 R1 Proc Natl Acad Sci USA 84:3038-3042(1987).
 RT "FUNCTION: CALDESMONIN IS A HEAT STABLE, ACTIN, CALMODULIN BINDING
 RT PROTEIN.
 RT "CA ALYIC ACTIVITY: ATR + protein + ATR + phosphoprotein.
 RT "SUBCELLULAR LOCATION: SUBSTANTIA: LOCALIZATION IN CERTAIN NEURONAL
 RT NUCLEI.
 RT "ALTERNATIVE PRODUCTS: CALCIUM/CALMODULIN DEPENDENT PROTEIN KINASE
 RT TYPE IV CATALYTIC CHAIN AND CALDESMONIN ARE OBTAINED BY ALTERNATIVE
 RT SPLICING.
 RT "TISSUE SPECIFICITY: BRAIN AND TESTIS.
 RT "SIMILARITY: RELATED TO THE GSK/HRP FAMILY OF PROTEIN KINASES.
 RT "FAM: SUBFAMILY.

CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).

DR EMBL: M16206; AAA19933.1;
 DR EMBL: M64266; AAA17364.1;
 DR EMBL: J03057; AAA17365.1;
 DR EMBL: X58995; CAA11741.1;
 DR PIR: A29878; A29878;
 DR PIR: S17656; S17656;
 DR BSSP: Q64450; 1A06;
 DR MGD: MGI:88258; Camk4;
 DR InterPro: IPR000719; Pak_kinase;
 DR InterPro: IPR002290; Ser_Thr_kinase;
 DR Pfam: PF00069; pkinase; 1;
 DR SMART: SM00220; S_PKc; 1;
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1;
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1;
 DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1;
 KW transferase; Serine/threonine-protein kinase; ATP binding;
 KW Calmodulin-binding; Alternative splicing;
 FT CHAIN 406 469 CALDESMONIN.
 FT DOMAIN 42 296 PROTEIN KINASE.
 FT NE_BIND 48 56 ATP (BY SIMILARITY).
 FT BINDING 71 71 ATP (BY SIMILARITY).
 FT ACT_SITE 160 160 BY SIMILARITY.
 FT DOMAIN 318 337 CALMODULIN BINDING (POTENTIAL).
 FT CONFLICT 278 280 VLD -> CEG1 (IN REF. 2).
 FT CONFLICT 302 302 N -> T (IN REF. 2).
 SQ SEQUENCE 469 AA: 52627 MW: 981986708221975 (K664)

Query Match 1.7% Score 9: DB 1: Length 469;
 Best Local Similarity 100.0% Pred. No. 0.47;
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 161 KIADEGLSK 169

DB 178 KIADEGLSK 186

RESULT 3

K0C4 HUMAN

ID K0C4 HUMAN STANDARD; PRI: 473 AA;
 AC Q16566;
 DT 01-NOV-1997 (Rel. 35, Created);
 DI 16-NOV-2001 (Rel. 40, Last sequence update);
 DE Calcium/calmodulin dependent protein kinase type IV catalytic chain
 (C47311.1.1.0) (CAM kinase-GK) (CaMK IV) [Contains: Caldesmonin].
 GN CAMK4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 9425404; PubMed 8089075;
 RA Kitani T., Okuno S., Fujisawa H.;
 RT "cDNA cloning and expression of human calmodulin dependent protein
 R1 kinase IV";
 R1 J. Biochem. 115:647-640(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE Cerebellum, and Thymus;
 RX MEDLINE 9425256; PubMed 8194751;
 RA Bland M.M., Monroe R.S., Ohmsted C.A.;
 RT "The cDNA sequence and characterization of the

RI Ca2+/calmodulin-dependent protein kinase α from human brain and
 RI thymus.";
 RI Gene 142:191-197(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RF TISSUE: Brain;
 RX MEDLINE 9114982; PubMed 8107230.
 RA Mosialos G., Hanissian S.H., Jawahar S., Vara L., Kieff E.,
 RA Chatila T.A.;
 RI "A Ca2+/calmodulin-dependent protein kinase (CAM kinase) expressed
 RI after transformation of primary human T-lymphocytes by Epstein Barr
 RI virus (EBV) is induced by the EBV oncogene LMP1";
 RL J. Virol. 68:1697-1705(1994).
 CC -!- FUNCTION: CALSPERMIN IS A HEAT-STABLE, ACIDIC, CALMODULIN-BINDING
 CC PROTEIN.
 CC -!- CATALYTIC ACTIVITY: ATP + protein -> ADP + O-phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: SUBSTANTIAL LOCALIZATION IN CERTAIN NEURONAL
 CC NUCLEI (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS: CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE
 CC TYPE IV CATALYTIC CHAIN AND CALSPERMIN ARE OBTAINED BY ALTERNATIVE
 CC SPLICING.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CAMK SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre/>
 CC or send an email to license@sib.ch).
 CC -----
 DR EMBL: D30742; BAA06463.1; ...
 DR EMBL: L17900; AAA35639.1; ...
 DR EMBL: L24959; AAA18251.1; ...
 DR HSSP: 053450; 1A06.
 DR MIM: 114080; ...
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR022930; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR SMART: SM00229; c_kinase.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_SF; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW Transmembrane, Serine/threonine protein kinase, Ca^{2+} binding,
 KW Calmodulin-binding; Alternative splicing.
 FT CHAIN 310 473 CALSPERMIN.
 FT DOMAIN 46 300 PROTEIN_KINASE.
 FT NP_BIND 52 60 ATP (BY SIMILARITY).
 FT BINDING 75 75 ATP (BY SIMILARITY).
 FT ACT_SITE 164 164 HY SIMILARITY.
 FT DOMAIN 322 341 CALMODULIN_BINDING (POTENTIAL).
 SQ SEQUENCE 473 AA; 51925 MW; EFER51E56 2426DC CRC64;

Query Match 1.7%; Score 9; Dc1; Length 473;
 Best Local Similarity 100.0%; Pred. No. 0.47;
 Matches 0; Conservation 0; Mismatches 0; Indels 0; Gaps 0

QY 161 KIAEGLSK 169
 11 11111
 DE 162 KIAEGLSK 169

RESULT 4

KCC4_RA1

ID KCC4_RA1 STANDARD; PRI: 474 AA.

AC P1234;

DI 01-JAN-1996 (Rel. 13, Created)

DI 01-APR-1993 (Rel. 25, Last sequence update)

DI 16-OCT-2001 (Rel. 40, Last annotation update)

DE Calcium/calmodulin-dependent protein kinase type IV catalytic chain

DE (EC 2.7.1.12) (CAM kinase α) (CAMK IV; Serotonia; Calspersmin).

GN CAMK4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciuroquathi; Muridae; Murinae; Rattus.
 OX NCBI:taxid:10116;
 RN [1]
 RI SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE 91288548; PubMed 1648240.
 RA Ohmstede C.-A., Riland M.M., Merrill B.M., Sahyoun N.;
 RI "Calcium-binding protein and Ca^{2+} /calmodulin-dependent protein
 RI kinase α and calspersmin: a gene within a gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5784-5788(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 91304083; PubMed 1649385.
 RA Means A.R., Cruzaleoni E., Iamagueresse B., Needleman P.S.,
 RA Slaughter G.R., Ono T.;
 RI "A novel Ca^{2+} /calmodulin-dependent protein kinase and a male germ
 RI cell-specific calmodulin-binding protein are derived from the same
 RI gene";
 RL Mol. Cell. Biol. 11:3960-3971(1991).
 RN [3]
 RP SEQUENCE OF 250-474 FROM N.A.
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE: Brain;
 RX MEDLINE 8917447; PubMed 2508431.
 RA Ohmsted C.-A., Jensen K.F., Sahyoun N.;
 RI "Ca2+/calmodulin-dependent protein kinase enriched in cerebellar
 RI granule cells. Identification of a novel neuronal
 RI calmodulin-dependent protein kinase";
 RL J. Biol. Chem. 264:2987-2992(1989).
 RN [4]
 RP SEQUENCE OF 335-474 FROM N.A., AND SEQUENCE OF 335-361.
 RC STRAIN-SPRAGUE-DAWLEY;
 RX MEDLINE 8912572; PubMed 2914893.
 RA Ono T., Slaughter G.R., Cook K.G., Means A.R.;
 RI "Molecular cloning sequence and distribution of rat calspersmin, a
 RI high affinity calmodulin-binding protein";
 RL J. Biol. Chem. 264:2981-2987(1989).
 CC -!- TISSUE SPECIFICITY: CAMK4 IS A TISSUE SPECIFIC PROTEIN KINASE,
 CC ENRICHED IN CEREBELLAR GRANULE CELLS.
 CC -!- FUNCTION: CALSPERMIN IS A HEAT-STABLE, ACIDIC, CALMODULIN-BINDING
 CC PROTEIN.
 CC -!- CATALYTIC ACTIVITY: ATP + protein -> ADP + O-phosphoprotein.
 CC -!- ALTERNATIVE PRODUCTS: CALCIUM/CALMODULIN-DEPENDENT PROTEIN KINASE
 CC TYPE IV CATALYTIC CHAIN AND CALSPERMIN ARE OBTAINED BY ALTERNATIVE
 CC SPLICING.
 CC -!- TISSUE SPECIFICITY: CALSPERMIN IS PREDOMINANTLY FOUND IN MAMMALIAN
 CC TESTIS AND BRAIN.
 CC -!- PTM: THE N-TERMINAL OF CALSPERMIN IS BLOCKED.
 CC -!- CAMK4 BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC CAMK SUBFAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre/>
 CC or send an email to license@sib.ch).
 CC -----

DR EMBL: M63344; AAA40865.1; ...
 DR EMBL: M74488; AAA40845.1; ALL_SEQ.
 DR EMBL: M64757; AAA40856.1; ...
 DR EMBL: M64757; AAA40872.1; ...
 DR EMBL: J04600; AAA11867.1; ...
 DR EMBL: J04446; AAA40890.1; ...
 DR PIR: A41103; TVRTP4.
 DR HSSP: P00518; 1PHK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR022930; Ser_thr_kinase.
 DR Pfam: PF00069; pkinase.1.
 DR SMART: SM00229; c_kinase.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

PF	PF00116	PROTEIN KINASE DOM 1
PF	PF00011	PROTEIN KINASE DOM 1
KW	3	3
KW	3	3
E	CHAIN	CHAIN
E1	DEMAN	DEMAN
E1	NEBND	NEBND
E1	BUNDNG	BUNDNG
E1	ACTGTT	ACTGTT
E1	DEMAN	DEMAN
E1	DEMAN	DEMAN
E1	DEMAN	DEMAN
E1	CHNLIC	CHNLIC
SE	SEQUENCE	SEQUENCE

Query Match: 1/100; Score: 9; DB: 1; Length: 474;
 Best Local Similarity: 100.00%; Prod. No.: 0.32;
 Matched: 0; Substituted: 0; Mismatches: 0; Indels: 0; Gaps: 0;

20	1961 KLAUF NISB	1966
21	1961 KLAUF NISB	1966
22	1961 KLAUF NISB	1966

SEQUENCE INFORMATION

KL076981 ARABID

LOCUS KL076981 STANFORD; REFSeq: Atg10. AA.

AT 200007; 20007; UniProt;

EF 15 JUL 2000 (Feb. 08, created)

EF 15 JUL 2000 (Feb. 08, last sequence update)

EF 05 MAR 2002 (Feb. 41, last annotation update)

GN SNF1 related protein kinase KIN10 (E1-2.2.1.) (AKIN10).

GN KIN10 OR AKIN10 OR A1G01090 OR IAP13.22.

OR Arabidopsis thaliana (Mouse ear cress).

OT Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:

OC eurosids II: Brassicales: Brassicaceae: Arabidopsis.

OX NCBI TaxId 3722.

EN 12

RF SEQUENCE FROM N.A.

RF MEDLINE 1611341; PubMed 109073.

EI Le Guen L., Thomas M., Bianchi M., Bailford N.G., Kreis M.:
"Structure and expression of a gene from Arabidopsis thaliana
encoding a protein related to SNF1 protein kinase.";
Gene 1. 1999. 254(1992).

SN 13

RF SEQUENCE FROM N.A.

RF STRAIN IV. COMMEDIA;

EI Legrand P., Kreis M., Thomas M.
Submitted (Oct 1999) to the EMBL/GenBank/CCDB Databases.

EN 13

RF SEQUENCE FROM N.A.

RF STRAIN IV. COMMEDIA;

FX MEDLINE 2010720; PubMed 1110713;

EA Schumacher M., Lemcke K., Klemer M., Amseure W., Hensel M.,
FA Erdmann R., Viller G., Blocker B., Perez Alonso M., Obermaier B.,
FA Berony H., Brady H., Davis H.L.A., Mahe K., Paradisemech F.,
FA De Simone V., Thiele M., Artigianze F., Robert C., Bretter F.,
FA Wincker P., Cottalero J., Weissenbach J., Saurin W., Quetier F.,
FA Schuster M., Muller Auer S., Edel B., Buchs M., Benes V.,
FA Weinbach J., Przenek H., Erle H., Jordan N., Banert S.,
FA Wedderhann K., Kraus H., Voss H., Holland R., Brandt P., Nyakatura G.,
FA Verzi A., DiAngelo M., Pallavicini A., Toppi S., Simionati B.,
FA Conrad A., Berninger K., Kauer M., Loehner T. H., Nordtsick G.,
FA Reicherz A., Scharte M., Scheon G., Barnes M., Terol J., Clement J.,
FA Navarro E., Clado P., Perez Perez A., Ottenwaelder B., Duchemin D.,
FA Cooke R., Landre M., Berger Laura G., Burnelle B., Masny D.,
FA de Bari M., Marise A.C., Albert J.P., Chetot A., Casanberta E.,
FA Martori A., Altirio A., Flores M., Espino R., Vitale D.,
FA Minhaup G., Boase D., Schoof B., Knoll S., Zaccaria P., Mewes H.W.,
FA Mayer F.F.X., Karl S., Town O.D., Kez H.H., Ballou L.J., Jenkins J.,
FA Kenney K.M., Walts A., Stetterback L., Ellis C.Y., Shea L.P.,

```

RA Cressy L.H., Haas B.J., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pail G., Militischer J., Sellers P., Gill J.E., Feldblum T.V.,
RA Pruss D., Lin X., Nicmar W.C., Salazar S.L., White G., Venter J.C.,
RA Fraser M., Kuroki T., Nakamura Y., Sato S., Kato T., Asami E.,
RA Sasamoto S., Kimura T., Idesawa Y., Kawashima K., Kishida Y.,
RA Kiyokawa T., Kohira M., Matsumoto M., Matsun A., Muraki A.,
RA Nakayama S., Nakazaki N., Shimizu M., Takeuchi O., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.:
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [4]
RP SEQUENCE OF 1 19 FROM N.A.
RC STRAIN:CV, COLUMBIA;
RX MEDLINE 95115691, Pubmed 7816049;
RA re Chen L., Thomas M., Kreis M.;
RT "Gene density and organization in a small region of the Arabidopsis
RL thaliana genome.";
RL Mol. Gen. Genet., 245:390-396(1994).
CC 1. FUNCTION. MAY PLAY AN IMPORTANT ROLE IN A SIGNAL TRANSDUCTION
CC CASCADE REGULATING GENE EXPRESSION AND CARBOHYDRATE METABOLISM IN
CC HIGHER PLANTS.
CC 2. TISSUE SPECIFICITY. EXPRESSED IN ROOTS, STEMS AND LEAVES.
CC 3. SIMILARITY. BELONGS TO THE SEP/TE FAMILY OF PROTEIN KINASES.
CC SNF1 SUBFAMILY.
CC
CC This SWISS Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
DR EMBL: M93023; AAA32736.1; -.
DR EMBL: X79707; CAA56146.1; -.
DR EMBL: A0098201; AAF26165.1; -.
DR EMBL: X94757; CAA64384.1; -.
DR HSSP: Q63450; 1A06.
DR Mendel: 14016; ARATH;2426;mn14016.
DR Mendel: 12898; ARATH;2098;1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001772; KAl.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR000449; UBA.
DR Pfam: PF02149; KAl; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00627; UBA; 1.
DR PRINTS: PR00109; TYRKINASE.
DR SMART: SM00320; C_Like; 1.
DR SMART: SM00165; UBA; 1.
DR ProSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR ProSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ProSITE: PS00011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP binding;
KW Phosphorylation.
FT DOMAIN 19 271 PROTEIN KINASE.
FT NF_BIND 25 33 ATP (BY SIMILARITY).
FT BINDING 48 48 ATP (BY SIMILARITY).
FT ACT_SITE 142 142 BY SIMILARITY.
FT MOD_RES 175 175 PHOSPHORYLATION (A16) (BY SIMILARITY).
SQ SEQUENCE 512 AA; 58373 MW; 5A18655A0AA508DF16c64;

```

Query Match: 1.7%; Score 9; DB 1; Length 512;
 Best Local Similarity: 100.0%; Pred. No. 0.4;
 Matches: 9; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

9y 160 VKIADFGLS 168
 |||||
 10b 156 VKIADFGLS 164


```

RESULT      8
SNF1_CANTR
25  SNF1_CANTR      STANHAKE,      PRI,      615 AA.
AC      094198:
DT      15-JUL-1999 (Ref. 48, Created)
DT      15-JUL-1999 (Ref. 48, Last sequence update)
DI      15-JUL-1999 (Ref. 48, Last annotation update)
DE      Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
GN      SNF1.
OS      Candida tropicalis (Yeast).
OC      Eukaryota, Eumyc, Ascomycota, Saccharomycotina, Saccharomycetes,
OC      Saccharomycetales, mitosporic Saccharomycetales; Candida.
OX      NCBI taxID 5482:
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Katali T., Nakawa K., Noda M., Tanaka A.:
RT      "Genetic evaluation of the function of SNF1 in Candida tropicalis."
RI      Submitted (EMBL) to the EMBL/GenBank/DBCLS databases.
Z1      1. FUNCTION. ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. II
Z2      INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
Z3      PROTEIN SNF4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
Z4      1. SUBCELLULAR LOCATION. ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY

```

```

RESULTS 19
SNF1_YEAST
ID SNF1_YEAST STANDARD: PRI: 633 AA.
AC P06782;
DI 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DI 15-JUL-1999 (Rel. 06, Last annotation update)
DE Carbon catabolite derepressing protein kinase (EC 2.7.1.1).
GN SNF1 OR CAT1 OR PCK1 OR PAS14 OR GLC2 OR YDR477W OR P8045.20.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Eumal; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces
OX NCBI_TaxID:4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE 86289463; PubMed 3526554;
RA Colenza J.L., Carlson M.;
RT "A yeast gene that is essential for release from glucose repression
RT encodes a protein kinase.";
RL Science 233:1175-1180(1986).
RF [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo K., Aviles E.,
RA Bruno A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunkeler-Smith S., Hyman R., Kemp C., Lashkar D., Lew H., Lin D.,
RA Mochlyais D., Nakahara K., Namath A., Ochner E., Oh C., Patel P.K.,

```

RA Roberts D., Schramm S., Schroeder M., Sjogren T., Shroff N.,
 RA Winant A., Yellon M., Feinstein D., Davis R.W.,
 RL Submitted (and 1995) to the EMBL/Genbank/DBPFI databases
 RN [4]
 RP SEQUENCE OF 274 281, 528 535 AND 622 633, AND PHOSPHORYLATION SITE.
 RX MEDLINE 9412398, PubMed 7505477.
 RA Mitchell K.L., Stapleton D., Cao G., House C., Mitchell P.,
 RA Katsis F., Witters L.A., Kemp B.E.,
 RT "Mammalian AMP-activated protein kinase shares structural and
 RT functional homology with the catalytic domain of yeast Snf1 protein
 RT kinase."
 RL J. Biol. Chem. 269:2361-2364(1994).
 CC -!- FUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT
 CC INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY
 CC PROTEIN SNF1. INTERACTS ALSO WITH SIPI1, SIPI2 AND GAL83. COULD
 CC PHOSPHORYLATES CAT8.
 CC -!- SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC SNF1 SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/faq.html>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M13971; AAA55058.1; -.
 DR EMBL: U33050; AAB64904.1; -.
 DR PIR: A26030; A26030.
 DR HSRP: U64450; 1A06.
 DR SIB: S000295; SNF1
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase: Serine/threonine protein kinase; ATP-binding;
 KW Phosphorylation; Carbohydrate metabolism; Nuclear protein.
 FT DOMAIN 18 32 POLY-HIS.
 FT DOMAIN 55 306 PROTEIN_KINASE.
 FT NP_BIND 61 69 ATP (BY SIMILARITY).
 FT BINDING 84 84 ATP (BY SIMILARITY).
 FT ACT_SITE 177 177 DY SIMILARITY.
 FT MOD_RES 210 210 PHOSPHORYLATION (AUTOPHOSPHORYLATION).
 SQ SEQUENCE 633 AA: 72045 MW: 1506356509664100 CR064.

Query Match 1.7% Score 9; DB 1; Length 633;
 Best local Similarity 100.0%; Pred. No. 6.49;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Gy 160 VKIADPGL 168
 |||||
 Db 191 VKIADPGL 199

RESULT 11

JAK3_HUMAN
 ID JAK3_HUMAN STANDARD; PRI: 1100 AA.
 AC 06/27/97
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tyrosine protein kinase JAK3 (E.C.2.7.1.12); Janus Kinase 3, (JAK 3).
 GN JAK3.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CC NCBI TaxID 10116;

RN [1]
 RP SEQUENCE FROM NIA.
 RC TISSUE: Spleen;
 RX MEDLINE 9412398, PubMed 7505477.
 RA Takahashi T., Shirasawa T.,
 RT "Molecular cloning of rat JAK3, a novel member of the JAK family of
 RT protein tyrosine kinases."
 RL FEBS Lett. 342:124-128(1994).
 CC -!- FUNCTION: TYROSINE KINASE OF THE NON-RECEPTOR TYPE. INVOLVED IN
 CC THE INTERLEUKIN 2 AND INTERLEUKIN 4 SIGNALING PATHWAY.
 CC PHOSPHORYLATES STAT6, IRS1, IRS2 AND PI3K.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: WHOLLY INTRACELLULAR, POSSIBLY MEMBRANE
 CC ASSOCIATED (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: TRANSCRIBED IN A VARIETY OF TISSUES INCLUDING
 CC SPLEEN, LUNG, KIDNEY AND INTESTINE.
 CC -!- DOMAIN: POSSESSES TWO PHOSPHOTRANSFERASE DOMAINS. THE SECOND ONE
 CC PROBABLY CONTAINS THE CATALYTIC DOMAIN (BY SIMILARITY). WHILE THE
 CC PRESENCE OF SLIGHT DIFFERENCES SUGGEST A DIFFERENT ROLE FOR
 CC DOMAIN 1.
 CC -!- PIM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2 AND IL-4 (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. JAK
 CC SUBFAMILY.
 CC -!- SIMILARITY: BELONGS TO THE SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/faq.html>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D28508; BAA05868.1; -.
 DR HSRP: P11462; 1A0W.
 DR InterPro: IPR000299; Band_4.1.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 2.
 DR SMART: SM00295; B41; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00219; Tykcs; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 2.
 DR PROSITE: PS00001; SH2; P458-NR01.
 KW Transferase: Tyrosine protein kinase; ATP-binding; Phosphorylation;
 KW SH2 domain; Repeat.
 FT DOMAIN 372 472 SH2 (ATYPICAL).
 FT DOMAIN 517 727 PROTEIN_KINASE_1.
 FT DOMAIN 818 1091 PROTEIN_KINASE_2.
 FT NP_BIND 824 832 ATP (BY SIMILARITY).
 FT BINDING 851 851 ATP (BY SIMILARITY).
 FT ACT_SITE 945 945 DY SIMILARITY.
 FT MOD_RES 974 976 PHOSPHORYLATION (AUTOPHOSPHORYLATION) (BY SIMILARITY).
 SQ SEQUENCE 1100 AA: 122560 MW: 10590A05F41E7EEL CR064.

Query Match 1.7% Score 9; DB 1; Length 1100;
 Best local Similarity 100.0%; Pred. No. 0.83;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Gy 159 BVKIADPGL 167
 |||||
 Db 250 BVKIADPGL 266

RESULT 12

JAK3_HUMAN
 ID JAK3_HUMAN STANDARD; PRI: 1124 AA

AC Q13359; Q13359; Q13360; Q13361;
 DI 1996 (Ref. 34) (created)
 DI 1996 (Ref. 34) (last sequence update)
 DI 1996 (Ref. 41) (last annotation update)
 DE Tyrosine protein kinase JAK3 (p145/J3.112) (Janus kinase 3) (JAK 3)
 DE (Janus-tyrosine kinase) (JAK3)
 GN JAK3
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
 NX NCBI TaxID: 9606;
 RN 1;
 RF SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE 9424381; PubMed 932770;
 RA Kikuchi M., McVicar D.W., Johnston J.A., Blake T.B., Chen Y.,
 RA Salter K.L., O'Shea J.A., Kelvin D.J., Staples J.E., Cortaldo D.R.,
 RA Chen Y.
 RT "Molecular cloning of JAK3, a Janus family protein tyrosine kinase
 RT expressed in natural killer cells and activated leukocytes."
 RL Proc. Natl. Acad. Sci. U.S.A. 91:6374-6378(1994).
 RN 3;
 RF SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
 RX MEDLINE 9607665; PubMed 759000;
 RA Lee K.S., Jan Y., Graham G.K., Withuhn R.A., Chiu J.K., Liu K.Y.,
 RT "A kinase-deficient splice variant of the human JAK3 is expressed in
 RT hematopoietic and epithelial cancer cells."
 RL J. Biol. Chem. 270:25028-25036(1995).
 RN 4;
 RF SEQUENCE OF 9191 FROM N.A.
 RX MEDLINE 9627845; PubMed 869278;
 RA Verbsky J.W., Bach E.A., Fung Y.F., Yang L., Randolph D.A.,
 RA Fields J.E.
 RT "Expression of a kinase-deficient human endothelial and other non-
 RT lymphoid and non-myeloid cells."
 RL J. Biol. Chem. 271:16976-16980(1996).
 RN 5;
 RF VARIANT WITH CYS 100.
 RX MEDLINE 956412; PubMed 765055;
 RA Manenti P., Villa A., Gilliani S., Sacco M.L., Fratini A., Porta F.,
 RA Gherli A., Johnston J.A., Candotti F., O'Shea J.J., Vezzoni P.,
 RA Notarangelo L.D.
 RT "Mutational inactivation of the human JAK3 gene in patients with natural severe combined
 RT immune deficiency (SCID)."
 RL Nature 377:69-72(1995).
 RN 6;
 RF VARIANT WITH GLY 481; 509; LEU 592 DEL AND ARG 759.
 RX MEDLINE 9992793; PubMed 954908;
 RA Candotti F., Stokes S.A., Johnston J.A., Gilliani S., Schumacher R.F.,
 RA Mellia P., Fiorini M., Ugueto A., Badolato R., Notarangelo L.D.,
 RA Bozzi F., Manenti P., Striano I., Vezzoni P., Blaese R.M., O'Shea J.J.,
 RA Villa A.
 RT "Structural and functional basis for JAK3 deficient severe combined
 RT immunodeficiency."
 RL Blood 85:996-1003(1997).
 RN 7;
 RF VARIANT WITH LEU 592.
 RX MEDLINE 9942994; PubMed 975972;
 RA Bozzi F., Torres G., Villa A., Badolato R., Schumacher R.F.,
 RA Khalil L., Casaleto J., Bresnani S., O'Shea J.J., Vezzoni P.,
 RA Notarangelo L.D., Candotti F.
 RT "Molecular and biochemical characterization of JAK3 deficiency in a
 RT patient with severe combined immunodeficiency over 20 years after
 RT bone marrow transplantation: implications for treatment."
 RL Br. J. Haematol. 102:1493-1496(1998).
 RN 8;
 RF VARIANT WITH ARG 141; THR 722 AND SER 910.
 RX MEDLINE 2014964; PubMed 1898195;
 RA Schumacher R.F., Mellia P., Badolato R., Fiorini M., Savoldi G.,
 RA Gilliani S., Villa A., Candotti F., Lampalini A., O'Shea J.J.,
 RA Notarangelo L.D.
 RT "Complete spontaneous reorganization of the human JAK3 gene and mutation
 RT analysis in severe combined immunodeficiency by single strand
 RT conformation polymorphism."

RL Hum. Genet. 106:73-79(2000).
 CC 1. FUNCTIONAL TYROSINE KINASE OF THE NON-RECEPTOR TYPE, INVOLVED IN
 CC THE INTERLEUKIN 2 AND INTERLEUKIN 4 SIGNALING PATHWAY.
 CC PHOSPHORYLATES STAT6, IRS1, IRS2 AND PI3K.
 CC 2. CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC 3. SUBCELLULAR LOCATION: WHOLLY INTRACELLULAR, POSSIBLY MEMBRANE
 CC ASSOCIATED (BY SIMILARITY).
 CC 4. ALTERNATIVE SPLICING: 2 ISOFORMS, 1/JAK3/1 AND 3/JAK3/
 CC 2/JAK3/2. 3/JAK3/2 (JAK3) (short form) and 1/JAK3/1 (JAK3)
 CC (JAK3) are produced by alternative splicing. Isoform 1 may be
 CC defective as it lack some part of the kinase domain.
 CC 5. TISSUE SPECIFICITY: IN NK CELLS AND AN NK-LIKE CELL LINE BUT NOT
 CC IN KILLING T CELLS OR IN OTHER TISSUES. THE 3 FORM IS MORE
 CC COMMONLY SEEN IN HEMATOPOIETIC LINES, WHEREAS THE 2 ARE 3 FORMS
 CC ARE DETECTED IN CELLS BOTH OF HEMATOPOIETIC AND EPITHELIAL
 CC ORIGINS.
 CC 6. DOMAIN: POSSESSES TWO TRANSFERKINASE DOMAINS. THE SECOND ONE
 CC PROBABLY CONTAINS THE CATALYTIC DOMAIN (BY SIMILARITY), WHILE THE
 CC PRESENCE OF SLIGHT DIFFERENCES SUGGEST A DIFFERENT ROLE FOR
 CC DOMAIN 1.
 CC 7. PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2 AND IL-4.
 CC 8. DISEASE: DEFECTS IN JAK3 ARE A CAUSE OF RECESSIVE T-CELL
 CC NEGATIVE/B CELL POSITIVE SEVERE COMBINED IMMUNODEFICIENCY (1-B/
 CC SCID). A CONDITION CHARACTERIZED BY THE ABSENCE OF CIRCULATING
 CC MATURE T-LYMPHOCYTES AND NK CELLS, NORMAL TO ELEVATED NUMBERS OF
 CC NONFUNCTIONAL B-LYMPHOCYTES, AND MARKED HYPOPLASIA OF LYMPHOID
 CC TISSUES.
 CC 9. SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES, JAK
 CC SUBFAMILY.
 CC 10. SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre/>
 CC or send an email to license@ebi.ac.uk).
 CC -----

DR EMBL: 009607; AAA19626.1; .
 DR EMBL: 041601; AAC50226.1; .
 DR EMBL: 041602; AAC50227.1; .
 DR EMBL: 057096; AAC50542.1; .
 DR HSSP: P11362; 1FGK.
 DR MIM: 600173; .
 DR MIM: 600802; .
 DR InterPro: IPR000299; Rand_4.1.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR01245; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 2.
 DR SMART: SM00295; B41; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00219; TykK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 2.
 DR PROSITE: PS50001; SH2; FALSE_NEG.
 KW Transferase; Tyrosine-protein kinase; ATP binding; Phosphorylation;
 KW SH2 domain; Repeat; Alternative splicing; Disease mutation; SCID.
 FT DOMAIN 375 475 SH2 (ATYPICAL).
 FT DOMAIN 521 781 PROTEIN KINASE 1.
 FT DOMAIN 822 1111 PROTEIN KINASE 2.
 FT RM_BI: 828 836 ATP (BY SIMILARITY).
 FT BINDING 855 855 ATP (BY SIMILARITY).
 FT ACT_SITE 949 949 BY SIMILARITY.
 FT MOD_RES 980 980 PHOSPHORYLATION (AUTO) (BY SIMILARITY).
 FT VARSPIC 1071 1124 HELMKLWAPSPQRRPSSSADLPLMLMSQSGPFIHAFI
 FT AHPCKRHSLSES - - SAAGASVSGSVIWAIVSVSKPAGA
 FT (IN ISOFORM 1).
 FT VARSPIC 1071 1124 HELMKLWAPSPQRRPSSSADLPLMLMSQSGPFIHAFI
 FT AHPCKRHSLSES - - SCYSWRDDLCSCGWPTVSRW

FI LAQSPQPRFLITATIVQLFFTLHATAASVAVENKIC (IN
 FT ISOFORM 2)
 FI VARIANT 100 100 Y -> C (IN SCID).
 FT /FTID-VAR_010424.
 FI VARIANT 151 151 D -> K (IN SCID).
 FT /FTID-VAR_010492.
 FI VARIANT 481 481 E -> G (IN SCID).
 FT /FTID-VAR_010493.
 FI VARIANT 582 582 K -> W (IN SCID).
 FT /FTID-VAR_010444.
 FI VARIANT 586 592 MISSING (IN SCID, LACK OF PHOSPHORYLATION
 FT IN RESPONSE TO CYTOKINE STIMULATION).
 FI /FTID-VAR_010495.
 FI VARIANT 722 722 V -> I (IN SCID).
 FT /FTID-VAR_010496.
 FI VARIANT 759 759 N -> K (IN SCID, CONSTITUTIVE
 FT PHOSPHORYLATION).
 FT /FTID-VAR_010497.
 FI VARIANT 910 910 L -> S (IN SCID).
 FT /FTID-VAR_010498.
 FI CONFLICT 34 34 A -> G (IN REF. 2).
 FI CONFLICT 117 117 MISSING (IN REF. 3).
 FI CONFLICT 187 187 I -> A (IN REF. 2).
 FI CONFLICT 212 212 A -> R (IN REF. 2).
 FI CONFLICT 222 222 E -> R (IN REF. 2).
 FI CONFLICT 610 610 M -> I (IN REF. 2).
 FI CONFLICT 845 846 AH -> GD (IN REF. 2).
 FI CONFLICT 896 897 PE -> QS (IN REF. 2).
 SQ SEQUENCE 1124 AA; 125015 MW; 1D9FD22368E0B8E4 CRC64;

Query Match: 17%; Score: 6; Gap: 1; Length 1124

Best local Similarity: 100.0%; Pred. No.: 84;

Matches: 9; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

Qy 159 HVKIADEGL 167
 IIIIIII
 DE 962 HVKIADEGL 970

RESULI 13
 JAK3_MOUSE
 ID JAK3_MOUSE STANDARD; PRT; 1299 AA
 AC Q62137; Q61747; Q61746;
 DE 01-NOV-1997 (Rel. 35, Created)
 DI 01-NOV-1997 (Rel. 35, Last sequence update)
 DI 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tyrosine-protein kinase JAK3 (EC 2.7.1.2) (Janus kinase 3) (JAK-3).
 GN JAK3
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
 CX NCBI TaxID: 10090
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE 9432920; PubMed 7518579.
 RA Pano S.G., Reddy E.P.;
 RT "JAK3: a novel JAK kinase associated with terminal differentiation of
 RT hematopoietic cells."
 RL Oncogene 9:2415-2423(1994).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX STRAIN PA16.0; X 12-82; FISHBY Thymus.
 RX MEDLINE 96184772; PubMed 8605329;
 PA Gurniak C.P., Berg L.J.;
 RT "Murine JAK3 is preferentially expressed in hematopoietic tissues and
 RT lymphocyte precursor cells."
 RL Blood 87:3151-3160(1996).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RX STRAIN PA16.0;
 RX MEDLINE 94291024; PubMed 8922486;
 RA Witthuhn B.A., Silvenhoinen O., Miura O., Lai K.S., Cwik C., Liu E.T.,

RA 1116 J.N.;
 RT "Isolation of the JAK3 gene kinase is regulated by Interferons
 RT 2 and 4 in lymphoid and myeloid cells."
 RL Nature 370:154-157(1994).

CC 1 FUNCTION: TYROSINE KINASE OF THE NON-RECEPTOR TYPE, INVOLVED IN
 CC THE INTERFERON-2 AND INTERLEUKIN-3 SIGNALING PATHWAY.
 CC PHOSPHORYLATES STAT5, IRS1, IRS2 AND FAK.
 CC 1-1 CATALYTIC ACTIVITY: ATP + a protein tyrosine + ADP -> protein
 CC tyrosine phosphate.
 CC 1-2 SUBCELLULAR LOCATION: WHOLLY INTRACELLULAR, POSSIBLY MEMBRANE
 CC ASSOCIATED (BY SIMILARITY).
 CC 1 ALTERNATIVE SPLICING: 3 EXONS, 1 (DOWN REK), 2 AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC 1 TISSUE SPECIFICITY: HIGHEST LEVELS FOUND IN THE THYMUS WITH
 CC SOMEWHAT LOWER LEVELS IN BONE MARROW, SPLEEN, HILAL LIVER AND
 CC ADULT CD4+ CD8- THYMOCYTES. VERY LOW LEVELS IN ADULT KIDNEY, LUNG,
 CC LIVER, BRAIN AND LIVER.
 CC 1 DOMAIN: POSSIBLY TWO PROTEIN KINASE DOMAINS, THE SECOND ONE
 CC PROBABLY CONTAINS THE CATALYTIC DOMAIN (BY SIMILARITY), WHILE THE
 CC PRESENCE OF SLIGHT DIFFERENCES SUGGEST A DIFFERENT ROLE FOR
 CC DOMAIN 1.
 CC 1 TIM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IL-2 AND IL-4.
 CC 1 SIMILARITY: PART OF THE JAK FAMILY OF PROTEIN KINASES, JAK
 CC SUBFAMILY.
 CC 1 SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see http://www.ebi.ac.uk/submit/).
 CC or send an email to license@ebi.ac.uk.

CC Swiss-Prot release 40, 1999-09-01
 DR EMBL: L33768; AAA21415.1;
 DR EMBL: L40172; AA142085.1;
 DR EMBL: L32955; AAA21565.1;
 DR BSSP: P11362; IAGW.
 DR MGI: MGI:99928; Jaks3.
 DR InterPro: IPR000299; Band 4.1.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001415; Tyr_pkinase.
 DR Pfam: PF00069; pkinase; 3.
 DR SMART: SM00295; H41; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_Dom; 2.
 DR PROSITE: PS50001; SH2; FALSE NEG.
 KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
 KW SH2 domain; Repeat; Alternative splicing.
 FT DOMAIN 548 588 SH2 (ATYPICAL).
 FT DOMAIN 732 992 PROTEIN KINASE 1.
 FT DOMAIN 1042 1295 PROTEIN KINASE 2.
 FT ND_BIND 1038 1046 ATP (BY SIMILARITY).
 FT BINDING 1064 1064 ATP (BY SIMILARITY).
 FT ACT_SITE 1158 1158 BY SIMILARITY.
 FT MOD_RES 1189 1189 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT VARSHLIC 222 228 PAWGFARPTATGLVY - LPSRLRSLRELLAAVNS
 FT (IN ISOFORM 3).
 FI VARSHLIC 222 228 VGLARITATAGLVY - EVVAALAEVY (IN
 FT ISOFORM 2).
 FI VARSHLIC 284 473 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FI VARSHLIC 491 498 QPTNSR - LAKRVEPAG (IN ISOFORM 2 AND
 FT ISOFORM 3).
 FI VARSHLIC 568 596 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 FI VARSHLIC 656 673 ASASPIEAGGSHALLLE - GLSQPHRLRELLAAVNS
 FT (IN ISOFORM 2 AND ISOFORM 3).
 FI VARSHLIC 1003 1027 LPHLAEVLIEMIAWPAALGIDSPRI - SDPTGPIPS
 FT FBEELVAGAGLYAGQDAIF (IN ISOFORM 2 AND
 FT ISOFORM 3).

```

E1 VARSPLE1 1299 1299 SPFTGGPSAP - EPHDKPAFAILSPDLPLWRGRPG
E1 (IN ISOPFORM 2 AND ISOPFORM 3).
E1 1. 98% IDENTITY & ACTIVITY.
E1 A - G (IN REF. 2).
E1 S - P (IN REF. 3).
E1 G - N (IN REF. 3).
E1 MISSING (IN REF. 3).
E1 K - N (IN REF. 3).
E1 A - P (IN REF. 2 AND 3).
E1 E - D (IN REF. 3).
E1 A - G (IN REF. 2 AND 3).
E1 N - Y (IN REF. 3).
E1 I - N (IN REF. 3).
E1 G - EW (IN REF. 2 AND 3).
E1 R - S (IN REF. 3).
E1 S - Q (IN REF. 3).
E1 30 QUENY 1299 AA 144114 MW: 155200.90AF(100%)664;

```

Query Match: 1.70k; Score 9; DB 1; Length 1299;
 Best Local Similarity: 100.00k; Prod. No. 0.71;
 Matches: 0; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

```

E2 1.10 HVFADEN 197
E2 1.11111
E2 1.11 HVFADEN 119

```

RESULT 11

```

CMB CHLOR
ID CMB_CHLOR STANDARD; PRT: 74 AA.
AC P09822;
DE 01-MAR-1989 (ref. 10, Created)
DE 01-MAR-1989 (ref. 10, Last sequence update)
DE 01-MAR-2002 (ref. 41, Last annotation update)
DE Ferredoxin-like protein.
GN FIXX OR RA0444 OR SMA0816.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID 382;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN 1021;
RX MEDLINE 87137267; PubMed 3029021;
RA Earl C.D., Ronson C.W., Ausubel F.M.;
RT "Genetic and structural analysis of the Rhizobium meliloti fixA,
DE fixB, fixC, and fixX genes.";
RL J. Bacteriol. 169:1127-1136(1987).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE 87165742; PubMed 3031610;
RA Iasha L., Kovalevskaya N., Samalov Z., Kondorosi A.;
RT "Rhizobium meliloti insertion element ISrm2 and its use for
DE identification of the fixX gene.";
RL J. Bacteriol. 169:1403-1409(1987).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN 1021;
RX MEDLINE 21396509; PubMed 11481432;
RA Barnett M.J., Fisher R.E., Jones T., Kemp C., Abela A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Gallibert F., Gouzy J.,
RA Gurjal M., Heng A., Hrizar L., Hyman F.W., Kahn D., Kahn M.H.,
RA Salhan S., Keating D.H., Pate C., Rock M.D., Surzycki R., Wells D.H.,
RA Yeh K.C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
DE Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9884-9888(2001).
CC 1. FUNCTION: COULD BE A 3E 4S CLUSTER CONTAINING PROTEIN.
CC 2. SIMILARITY: TO FERREDOXINS FROM PSEUDOMONAS PUTIDA AND PSEUDOTUM
DE TARTARIVIBRIUM, FERREDOXIN 1 FROM ACETIVIBRIO VINELANDII, FERREDOXIN
DE 11 FROM DESULFOVIBRIO DESULFURICANS.
CC 3. THIS SWISS PROT entry is copyright. It is produced through a collaboration
DE between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
DE use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
DE entities requires a license agreement (see http://www.ebi.ac.uk/ebis/seq/oln/seq\_send\_email\_to\_license@ebi.ac.uk).
CC 4. EMBL: M15546; AAA21771.1;
DE PROTEIN: Ferredoxin-like bacteriochlorophyll; Repeat: Chlorosome.
E1 UN1 MF 1 0 BY SIMILARITY.
E1 DOMAIN 1 60 4 X APPROXIMATE REPEATS.

```

```

E1 REPEAT 7 16 1.
E1 REPEAT 28 38 2.
E1 REPEAT 39 49 3.
E1 REPEAT 50 60 4.
E1 SEQUENCE 74 AA: 7477 MW: F5692E9143E1670E CRC64;

```

Query Match: 1.50k; Score 8; DB 1; Length 74;
 Best Local Similarity: 100.00k; Prod. No. 0.71;
 Matches: 0; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

```

Oy 314 LQSVSSAI 321
Oy 11111111
Db 63 LQSVSSAI 70

```

RESULT 15

```

FIXX_RHIME
ID FIXX_RHIME STANDARD; PRT: 98 AA.
AC P09822;
DE 01-MAR-1989 (ref. 10, Created)
DE 01-MAR-1989 (ref. 10, Last sequence update)
DE 01-MAR-2002 (ref. 41, Last annotation update)
DE Ferredoxin-like protein.
GN FIXX OR RA0444 OR SMA0816.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID 382;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN 1021;
RX MEDLINE 87137267; PubMed 3029021;
RA Earl C.D., Ronson C.W., Ausubel F.M.;
RT "Genetic and structural analysis of the Rhizobium meliloti fixA,
DE fixB, fixC, and fixX genes.";
RL J. Bacteriol. 169:1127-1136(1987).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE 87165742; PubMed 3031610;
RA Iasha L., Kovalevskaya N., Samalov Z., Kondorosi A.;
RT "Rhizobium meliloti insertion element ISrm2 and its use for
DE identification of the fixX gene.";
RL J. Bacteriol. 169:1403-1409(1987).
RN 13
RP SEQUENCE FROM N.A.
RC STRAIN 1021;
RX MEDLINE 21396509; PubMed 11481432;
RA Barnett M.J., Fisher R.E., Jones T., Kemp C., Abela A.P.,
RA Barloy-Hubler F., Bowser L., Capela D., Gallibert F., Gouzy J.,
RA Gurjal M., Heng A., Hrizar L., Hyman F.W., Kahn D., Kahn M.H.,
RA Salhan S., Keating D.H., Pate C., Rock M.D., Surzycki R., Wells D.H.,
RA Yeh K.C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
DE Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9884-9888(2001).
CC 1. FUNCTION: COULD BE A 3E 4S CLUSTER CONTAINING PROTEIN.
CC 2. SIMILARITY: TO FERREDOXINS FROM PSEUDOMONAS PUTIDA AND PSEUDOTUM
DE TARTARIVIBRIUM, FERREDOXIN 1 FROM ACETIVIBRIO VINELANDII, FERREDOXIN
DE 11 FROM DESULFOVIBRIO DESULFURICANS.
CC 3. THIS SWISS PROT entry is copyright. It is produced through a collaboration
DE between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
DE use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
DE entities requires a license agreement (see http://www.ebi.ac.uk/ebis/seq/oln/seq\_send\_email\_to\_license@ebi.ac.uk).
CC 4. EMBL: M15546; AAA21771.1;
DE PROTEIN: Ferredoxin-like bacteriochlorophyll; Repeat: Chlorosome.
E1 UN1 MF 1 0 BY SIMILARITY.
E1 DOMAIN 1 60 4 X APPROXIMATE REPEATS.

```

DR PIR: D26952; D26952.
DR PIR: A26933; A26933.
DR InterPro: IPR001450; 4Fe4S_ferredoxin.
DR Pfam: PF0007; fer4; 1.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; 1.
KW Electron transport; Iron-sulfur; Nitrogen fixation; Plasmid;
KW Complete proteome.
FT CONFLICT 62 62 I -> L (IN REF. 2).
FT CONFLICT 97 97 E -> S (IN REF. 2).
SQ SEQUENCE 98 AA: 10937 MW: 0201C1A9254C2D8 CRC64;

Query Match 1.5%; Score 8; 18 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 0.93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 RSPSLNLL 535
| | | | | | | |
Db 33 RSPSLNLL 40

Search completed: May 14, 2002, 21:45:29
Job time: 329 sec



GenCore version: 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 21:22:46, Search time 52.98 seconds
(without alignments)
979 393 Million cell updates/sec

Title: US-09-445-223-1
Perfect score: 540
Sequence: 1 MNGEAICSALEPIPYHKLAD.....PEILVVSRSPLNLIAQKSM 540

Scoring table: Q1100
Gapop 60.0, Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: PIR_71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	10	1/9	517	2	T23413	hypothetical prote
2	9	1/7	52	2	S56723	SNF1-related prote
3	9	1/7	66	2	I49118	ink 76 - mouse (fr
4	9	1/7	67	2	I48695	gene ask1 protein
5	9	1/7	416	2	T29805	hypothetical prote
6	9	1/7	417	2	T33827	hypothetical prote
7	9	1/7	440	2	T33070	hypothetical prote
8	9	1/7	441	2	T29186	hypothetical prote
9	9	1/7	459	1	S17556	Ca2+/calmodulin-de
10	9	1/7	472	2	B90100	SNF1-related kinase
11	9	1/7	473	1	S59941	serine/threonine-s
12	9	1/7	473	1	A53036	Ca2+/calmodulin de
13	9	1/7	474	1	IVR104	Ca2+/calmodulin de
14	9	1/7	502	2	I52637	Ca2+/calmodulin-de
15	9	1/7	504	2	T19449	probable serine/th
16	9	1/7	511	1	A56009	serine/threonine-s
17	9	1/7	512	1	J01446	serine/threonine-s
18	9	1/7	512	2	I07789	probable serine/th
19	9	1/7	562	2	T29858	hypothetical prote
20	9	1/7	576	2	I41587	probable carbon de
21	9	1/7	582	2	S72517	serine/threonine-s
22	9	1/7	600	1	A27436	PKC2 protein - yea
23	9	1/7	863	2	F85343	hypothetical prote
24	9	1/7	871	1	I48695	protein-tyrosine k
25	9	1/7	881	1	I48697	protein-tyrosine k
26	9	1/7	892	2	T08973	probable serine/th
27	9	1/7	946	1	A47299	thor related recept
28	9	1/7	1099	2	S48053	protein tyrosine k
29	9	1/7	1100	2	S43677	protein tyrosine k

30	9	1/7	1124	2	A95747	L-JAK protein-tyro
31	9	1/7	1299	2	I58401	protein tyrosine k
32	8	1/5	43	2	I34711	kinase-related pro
33	8	1/5	49	2	I41615	G protein-coupled
34	8	1/5	50	2	I38226	protein-tyrosine k
35	8	1/5	52	2	S56722	SNF1-related prote
36	8	1/5	56	2	S66324	protein kinase AK1
37	8	1/5	68	1	PEKMX	terredoxin [4Fe-4S
38	8	1/5	98	2	I05317	FlxX terredoxin-1
39	8	1/5	101	2	S51021	protein kinase PKX
40	8	1/5	101	2	S51162	protein kinase PKX
41	8	1/5	112	2	A49550	AMP-activated prot
42	8	1/5	156	2	I12201	probable ede2 like
43	8	1/5	174	2	G01743	AMP-activated prot
44	8	1/5	281	2	P04220	protein kinase (E)
45	8	1/5	291	2	I45698	hypothetical prote

ALIGNMENTS

RESULT 1
I23413
hypothetical protein K07F5.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 17 Oct 1999 #sequence_revision 15 Oct 1999 #text_change 17 Mar 2000
C:Accession: I23413
R:Hembry, C.
submitted to the EMBL data library, March 1996
A:Reference number: Z19738
A:Accession: I23413
A:Title: C. elegans protein kinase (PK) family member, C. elegans
A:Molecule type: DNA
A:Residues: 1-517 -WHL-
A:Cross-references: EMBL:Z19738; PIR:K07F5.4; CDD:K07F5.4
A:Experimental source: clone K07F5
C:Genetics:
A:Gene: CESP:K07F5.4
A:Map position: 4
A:Strain: 4132 (P1) - P1 (P1) - P1 (P1) - P1 (P1)
C:Superfamily: protein tyrosine kinase ser: protein kinase homology: SH2 homology: SH

Query Match 1.9%; Score 10; DB 2; Length 517;
Best Local Similarity 100.0%; Pred. No. 0.16;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADPECLSK 169
| | | | | | | | | |
DB 296 VKIADPECLSK 295

RESULT 2
S56723
SNF1-related protein kinase (EC 2.7.1.) (clone AspK) - oat (fragment)
C:Species: Avena sativa (oat)
C:Date: 27 Oct 1995 #sequence_revision 19 Jan 1996 #text_change 10 Sep 1997
C:Accession: S56723
R:Huttly, A.K.; Phillips, A.L.
Plant Mol. Biol. 27, 1043-1052, 1995
A:Title: Gibberellin-regulated expression in oat aleurone cells of two kinases that s
A:Reference number: S56636; M11095284341
A:Accession: S56723
A:Molecule type: mRNA
A:Residues: 1-52 -SHU1-
C:Superfamily: phosphotransferase/Threonine Tyrosine Kinase; protein kinase hom
C:Keywords: phosphotransferase; protein kinase
P1: S2T323; protein kinase homology (fragment) - P1

Query Match 1.7%; Score 9; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 100 VK ADFGL 168
11 111
16 100 VK ADFGL 168

RESULT 1 6
T33827
hypotheical protein F23C8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29 Oct 1999 #sequence_revision: 25 Oct 1999 #text_change: 31 Jan 2000
C:Accession: T33827
R:Miller, N.; Wamsley, P.
submitted to the EMBL Data Library, November 1999
A:Description: The sequence of C. elegans cosmid F23C8.
A:Reference number: Z21419
A:Accession: T33827
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-416 (+LAT)
A:Cross references: EMBL:AF061958; FIDN:AA06183.1; RPTB:GN00019; CESP:C25A8.5
A:Experimental source: Strain Bristol N2; clone C25A8
C:Genetics:
A:Gene: CESP:C25A8.5
A:Map position: 4
A:Index: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
C:Superfamily: protein-tyrosine kinase src; protein kinase homology: SH2 homology: SH

Query Match 1.7%; Score 9; DB 2; Length 66;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 100 VK ADFGL 168
11 111
16 100 VK ADFGL 168

RESULT 1 4
T33827
hypotheical protein F23C8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29 Oct 1999 #sequence_revision: 25 Oct 1999 #text_change: 31 Jan 2000
C:Accession: T33827
R:Miller, N.; Wamsley, P.
submitted to the EMBL Data Library, November 1999
A:Description: The sequence of C. elegans cosmid F23C8.
A:Reference number: Z21419
A:Accession: T33827
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-416 (+LAT)
A:Cross references: EMBL:AF061958; FIDN:AA06183.1; RPTB:GN00019; CESP:C25A8.5
A:Experimental source: Strain Bristol N2; clone C25A8
C:Genetics:
A:Gene: CESP:C25A8.5
A:Map position: 4
A:Index: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
C:Superfamily: protein-tyrosine kinase src; protein kinase homology: SH2 homology: SH

Query Match 1.7%; Score 9; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 100 VK ADFGL 168
11 111
16 100 VK ADFGL 168

RESULT 1 5
T33827
hypotheical protein F23C8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29 Oct 1999 #sequence_revision: 25 Oct 1999 #text_change: 31 Jan 2000
C:Accession: T33827
R:Miller, N.; Wamsley, P.
submitted to the EMBL Data Library, November 1999
A:Description: The sequence of C. elegans cosmid F23C8.
A:Reference number: Z21419
A:Accession: T33827
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-416 (+LAT)
A:Cross references: EMBL:AF061958; FIDN:AA06183.1; RPTB:GN00019; CESP:C25A8.5
A:Experimental source: Strain Bristol N2; clone C25A8
C:Genetics:
A:Gene: CESP:C25A8.5
A:Map position: 4
A:Index: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
C:Superfamily: protein-tyrosine kinase src; protein kinase homology: SH2 homology: SH

A:Reference number: Z20689
A:Accession: T29805
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-416 (+LAT)
A:Cross references: EMBL:AF061958; FIDN:AA06183.1; RPTB:GN00019; CESP:C25A8.5
A:Experimental source: Strain Bristol N2; clone C25A8
C:Genetics:
A:Gene: CESP:C25A8.5
A:Map position: 4
A:Index: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
C:Superfamily: protein-tyrosine kinase src; protein kinase homology: SH2 homology: SH

Query Match 1.7%; Score 9; DB 2; Length 416;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 VKIADFGLS 168
111111111
Db 258 VKIADFGLS 266

RESULT 1 6
T33827
hypotheical protein F23C8.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29 Oct 1999 #sequence_revision: 25 Oct 1999 #text_change: 31 Jan 2000
C:Accession: T33827
R:Miller, N.; Wamsley, P.
submitted to the EMBL Data Library, November 1999
A:Description: The sequence of C. elegans cosmid F23C8.
A:Reference number: Z21419
A:Accession: T33827
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-417 (+M1)
A:Cross references: EMBL:AF061958; FIDN:AA06183.1; RPTB:GN00019; CESP:C25A8.5
A:Experimental source: Strain Bristol N2; clone C25A8
C:Genetics:
A:Gene: CESP:C25A8.5
A:Map position: 4
A:Index: 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60,61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98,99,100
C:Superfamily: protein-tyrosine kinase src; protein kinase homology: SH2 homology: SH

Query Match 1.7%; Score 9; DB 2; Length 417;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 VKIADFGLS 168
111111111
Db 258 VKIADFGLS 266

RESULT 1 7
T33070
hypotheical protein C35E7.10a - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29 Oct 1999 #sequence_revision: 29 Oct 1999 #text_change: 04 Mar 2000
C:Accession: T33070
R:Graves, T.; McDonald, R.
submitted to the EMBL Data Library, May 1998
A:Description: The sequence of C. elegans cosmid C35E7.
A:Reference number: Z21278
A:Accession: T33070
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-440 (+GRA)
A:Cross references: EMBL:AF067216; FIDN:AA17525.1; RPTB:GN00019; CESP:C35E7.10a
A:Experimental source: Strain Bristol N2; clone C35E7
C:Genetics:
A:Gene: CESP:C35E7.10a

A:Map position: 1
 A:Residues: 129-261 (133%) (100%) (100%) (100%) (100%) (100%)
 C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3

Query Match 1.7% Score 9; DB 2; Length 430;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 VKIADFGLS 168
 111111111
 Db 261 VKIADFGLS 269

RESULT 8
 129186
 hypothetical protein C55C3.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Sep-1999 #sequence_revision 15-Sep-1999 #text_change 04-Mar-2000
 C:Accession: 129186
 R:Woessner, J.; Stellyes, L.
 submitted to the EMBL Data Library, April 1996
 A:Description: The sequence of C. elegans csmid C55C3.
 A:Reference number: Z20585
 A:Accession: 129186
 A:Status: preliminary; translated from CR/218/TPP1
 A:Molecule type: DNA
 A:Residues: 1-431 (100%)
 A:Cross-references: EMBL:Z20585; PDB:AAA99-60.1; GSPDB:GN02022; GSPDB:G5507.4
 A:Experimental source: strain Bristol N2; clone C55C3
 C:Genetics
 A:Gene: CESP:C55C3.4
 A:Map position: 4
 A:Residues: 129-261 (133%) (100%) (100%) (100%) (100%) (100%)
 C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3

Query Match 1.7% Score 9; DB 2; Length 431;
 Best Local Similarity 100.0%; Pred. No. 1.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 VKIADFGLS 168
 111111111
 Db 261 VKIADFGLS 269

RESULT 9
 S17656
 Ca2+/calmodulin-dependent protein kinase (EC 2.7.1.123) IV mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S17656; A29878; I49571
 R:Jones, D.A.; Glod, J.; Wilson-Shaw, D.; Hahn, W.E.; Sikela, J.M.
 FEBS Lett. 289, 105-109, 1991
 A:Title: cDNA sequence and differential expression of the mouse Ca(2+)/calmodulin-dependent
 A:Reference number: S17656; MUID:91372388
 A:Accession: S17656
 A:Molecule type: mRNA
 A:Residues: 1-469 (100%)
 A:Cross-references: EMBL:X58945; NID:q50366; PDB:AAA1741.1; PDB:q50367
 R:Sikela, J.M.; Hahn, W.E.
 Proc. Natl. Acad. Sci. U.S.A. 84, 3048-3042, 1987
 A:Title: Screening an expression library with a ligand probe: isolation and sequence of
 A:Reference number: A29878; MUID:97204263
 A:Accession: A29878
 A:Molecule type: mRNA
 A:Residues: 315-469 (100%)
 A:Cross-references: GB:M16206; NID:q200360; PDB:AAA39933.1; PDB:q387512
 R:Sikela, J.M.; Law, M.L.; Kan, P.
 Genomics 4, 21-27, 1989
 A:Title: Chromosomal localization of the human gene for brain Ca2+/calmodulin-dependent
 A:Reference number: I49571; MUID:89122027
 A:Accession: I49571

A:Status: preliminary; translated from CR/218/TPP1
 A:Molecule type: mRNA
 A:Residues: 129-261 (133%) (100%) (100%) (100%) (100%) (100%)
 A:Cross-references: GB:Z20585; NID:q192366; PDB:AAA97-60.1; PDB:q192367
 A:Experimental source: brain
 C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology
 C:Keywords: alternative splicing; ATP; calmodulin binding; phosphotransferase; serine
 P:48-256/sequence; protein kinase; h2.1.3; EINS
 P:48-56/sequence; protein kinase ATP-binding motif
 P:309-469/product; calipermin #status predicted (GSP)
 P:318-377/sequence; calmodulin binding #status predicted
 P:71/Active site: Lys #status predicted

Query Match 1.7% Score 9; DB 1; Length 469;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 161 KIAEPHLSK 169
 111111111
 Db 178 KIAEPHLSK 186

RESULT 10
 B90100
 GTP-related kinase (unrelated) - Galliardia theta nucleomorph
 C:Species: nucleomorph Galliardia theta
 A:Note: a nucleomorph in the vestigial nucleus of a eukaryotic endosymbiont
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-Jun-2001
 C:Accession: B90100
 R:Bouglas, S.; Zanner, P.; Brande, H.; Beaton, M.; Denay, S.; Denay, L.; Wu, X.;
 Nature 410, 1091-1096, 2001
 A:Title: The highly reduced genome of an enslaved alga nucleus.
 A:Residues: 1-472 (100%)
 A:Accession: B90100
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-472 (100%)
 A:Cross-references: CR:AF155818; NID:q13794554; PDB:AAK39929.1; GSPDB:GN09150
 C:Genetics
 A:Gene: kin(snfl)
 A:Map position: 1
 A:Genome: nucleomorph
 C:Superfamily: AMU-activated protein kinase; protein kinase homology
 C:Keywords: nucleomorph

Query Match 1.7% Score 9; DB 2; Length 472;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 VKIADFGLS 168
 111111111
 Db 149 VKIADFGLS 157

RESULT 11
 S59941
 serine/threonine-specific protein kinase (EC 2.7.1.1) BKIN2 barley (fragrant)
 C:Species: Hordeum vulgare (barley)
 C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 26-Feb-1999
 C:Accession: S59941
 R:Hannappel, U.; Vicente-Carvajal, J.; Parker, J.R.A.; Showry, P.K.; Halford, N.D.
 Plant Mol. Biol. 27, 1245-1249, 1995
 A:Title: Differential expression of two barley SNPL-related protein kinase genes
 A:Reference number: S59941; MUID:95284374
 A:Accession: S59941
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-473 (100%)
 A:Cross-references: EMBL:X52548
 C:Genetics
 A:Gene: BKIN2

F:48-57/Region: protein kinase ATP-binding motif
 F:95-474/Region: calmodulin #status predicted CSP
 F:48-337/Region: calmodulin binding #status predicted
 F:71/Active site: lys #status predicted

Query Match 1.7% Score 9; DB 1; Length 474;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 161 KIADEGLSK 169
 Db 178 KIADEGLSK 186

RESULT 14
 152637
 Ca2+/calmodulin dependent protein kinase IV beta polypeptide - rat
 C:Species: Rattus sp. (rat)
 C:Date: 26-Jul-1996 #sequence_revision 26 Jul 1996 #text_change 18-Jun-1999
 C:Accession: 152637
 R:Sakagami, H.; Kondo, H.
 Brain Res. Mol. Brain Res. 19, 215-218, 1993
 A:Title: Cloning and sequencing of a gene encoding the beta polypeptide of Ca2+/calmodul
 A:Reference number: 152637; MUID:94018484
 A:Accession: 152637
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-502 <RES>
 A:Cross-references: GB:S65840; NID:g425483; UDN:AAB29372.1; PID:g425384
 C:Superfamily: Ca2+/calmodulin-dependent protein kinase; protein kinase homology
 C:Keywords: ATP; calmodulin binding
 F:68-324/Region: protein kinase homology <KIN>
 F:76-84/Region: protein kinase ATP-binding motif

Query Match 1.7% Score 9; DB 2; Length 502;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 161 KIADEGLSK 169
 Db 206 KIADEGLSK 214

RESULT 15
 T16449
 probable serine/threonine-specific protein kinase (EC 2.7.1.-) - cucumber
 N:Alternate names: SNF1-related protein kinase
 C:Species: Cucumis sativus (cucumber)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
 C:Accession: T16449
 R:Gumpel, N.J.
 submitted to the EMBL Data Library, December 1996
 A:Reference number: Z17020
 A:Accession: T16449
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 1-504 <GUM>
 A:Cross-references: EMBL:Y10036
 A:Experimental source: cv. Masterpiece; cotyledon
 C:Function:
 A:Description: catalyzes the formation of peptidyl-serine phosphate or peptidyl-threonin
 C:Superfamily: AMF-activated protein kinase; protein kinase homology
 C:Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 F:6-260/Region: protein kinase homology <KIN>

Query Match 1.7% Score 9; DB 2; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 VKIADGLS 168

|||||||
 Db 145 VKIADGLS 153

Search completed: May 14, 2002, 21:34:44
 Job time: 658 sec

QY 160 VKIALFGLS 168
 11111111
 DB 16 VKIALFGLS 18

RESULT 5
 I49118
 link 76 - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1996
 C:Accession: I49118
 R:Siyanova, E.Y.; Serfas, M.S.; Mazo, I.A.; Tyner, A.L.
 Oncogene 9, 2053-2057, 1994
 A:Title: Tyrosine kinase gene expression in the mouse small intestine.
 A:Reference number: I49118; MUID:94268846
 A:Accession: I49118
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-66 <RES>
 A:Cross-references: EMBL:U13259; NID:532727; EMBL:AAB0447.1; EMBL:542728
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
 F01-66/Domain: protein kinase homology (fragment) (KIN)

Query Match 1.7%; Score 9; DB 1; Length 66;
 Best Local Similarity 100.0%; Pred. No. 1; 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 160 VKIADPGLS 168
 11111111
 DB 18 VKIADPGLS 26

RESULT 4
 I48695
 gene nsk1 protein - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1996
 C:Accession: I48695
 R:Keith, A.D.
 Mamm. Genome 6, 689, 1995
 A:Title: Isolation and chromosomal location of nsk1, a novel murine putative receptor ty
 A:Reference number: I48695; MUID:96081359
 A:Accession: I48695
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-67 <RES>
 A:Cross-references: EMBL:X86367; NID:41091665; EMBL:AAB0127.1; EMBL:1386
 C:Genetics:
 A:Gene: Nsk1
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
 F01-67/Domain: protein kinase homology (fragment) (KIN)

Query Match 1.7%; Score 9; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1; 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 160 VKIADPGLS 168
 11111111
 DB 17 VKIADPGLS 25

RESULT 5
 T29805
 hypothetical protein C25A8.5 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T29805
 R:Latreille, P.; Stellyes, L.
 Submitted to the EMBL Data Library, June 1996
 A:Description: The sequence of C. elegans cosmid C25A8.

A:Reference number: Z20689
 A:Accession: T29805
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-416 <IAI>
 A:Cross-references: EMBL:U13259; NID:532727; EMBL:AAB0447.1; EMBL:542728
 A:Experimental source: strain Bristol; N2; clone C25A8
 C:Genetics:
 A:Gene: CESP:C25A8.5
 A:Map position: 4
 A:Introns: 22/3; 60/3; 126/3; 66/3; 66/3
 C:Superfamily: protein tyrosine kinase domain protein kinase homology (KIN)

Query Match 1.7%; Score 9; DB 1; Length 416;
 Best Local Similarity 100.0%; Pred. No. 1; 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 160 VKIADPGLS 168
 11111111
 DB 258 VKIADPGLS 266

RESULT 6
 T33827
 hypothetical protein F338.7 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T33827
 R:Miller, N.; Wamsley, P.
 Submitted to the EMBL Data Library, November 1996
 A:Description: The sequence of C. elegans cosmid F338.7
 A:Reference number: Z214.9
 A:Accession: T33827
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-417 <MIL>
 A:Cross-references: EMBL:AF057583; EMBL:AAB0127.1; EMBL:U13259; EMBL:542728
 A:Experimental source: strain Bristol; N2; clone F338.7
 C:Genetics:
 A:Gene: CESP:F338.7
 A:Map position: 1
 A:Introns: 22/3; 60/3; 126/3; 66/3; 66/3
 C:Superfamily: protein tyrosine kinase domain protein kinase homology (KIN)

Query Match 1.7%; Score 9; DB 1; Length 417;
 Best Local Similarity 100.0%; Pred. No. 1; 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 160 VKIADPGLS 168
 11111111
 DB 258 VKIADPGLS 266

RESULT 7
 T33670
 hypothetical protein C3367.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Mar-2000
 C:Accession: T33670
 R:Graves, T.; McDonald, K.
 Submitted to the EMBL Data Library, May 1996
 A:Description: The sequence of C. elegans cosmid C3367.10
 A:Reference number: Z21.78
 A:Accession: T33670
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-430 <GRA>
 A:Cross-references: EMBL:AF057516; EMBL:AAB0127.1; EMBL:U13259; EMBL:542728
 A:Experimental source: strain Bristol; N2; clone C3367
 C:Genetics:
 A:Gene: CESP:C3367.10a



Query Match 36.1%; Score 195; DI 4; Length 264;
Best Local Similarity 100.0%; Pred. No. 4.e-189;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVWPLRFRILHEIALGVNYLHNMT 139
Db 70 PILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVWPLRFRILHEIALGVNYLHNMT 129
QY 140 PPLLHDLKTQNILLDNEFHVKTADFGLSKWRMMSLSOSRSSKSAPEGGTIIYMPPEYNE 199
Db 130 PPLLHDLKTQNILLDNEFHVKTADFGLSKWRMMSLSOSRSSKSAPEGGTIIYMPPEYNE 189
QY 200 PGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSGQHRPVINEESLPYDI 259
Db 190 PGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLQIMYSVSGQHRPVINEESLPYDI 249
QY 260 PHARMISLIESGWA 274
Db 250 PHARMISLIESGWA 264

RESULT 10
US-09-069-023-6
; Sequence 5, Application US/09069023A
; Patent No. 6348574
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Koseki, Takeyoshi
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; FILE REFERENCE: UM-03333
; CURRENT APPLICATION NUMBER: US/09/069,023A
; CURRENT FILING DATE: 1998-04-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 6
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-6

Query Match 30.9%; Score 167; DI 4; Length 167;
Best Local Similarity 100.0%; Pred. No. 5.3e-161;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 SRKAQDCYFMKLIHHCPCGNHSDSTISGSQRAAFCDHKITPCSSAIINPLSTAGNSERLQP 433
Db 1 SRKAQDCYFMKLIHHCPCGNHSDSTISGSQRAAFCDHKITPCSSAIINPLSTAGNSERLQP 60
QY 434 GIAQOWIQSKREIVNQMTAEACLNQSLDALLSRDLINKEDYELVSTKPTRTSKVRQLLDT 493
Db 61 GIAQOWIQSKREIVNQMTAEACLNQSLDALLSRDLINKEDYELVSTKPTRTSKVRQLLDT 120
QY 494 IDICGEEFAKVIVQKLNKQMQCLQPYPEILVVSRSISLNLLQNKSM 540
Db 121 IDICGEEFAKVIVQKLNKQMQCLQPYPEILVVSRSISLNLLQNKSM 167

RESULT 11
US-09-099-041A-5
; Sequence 5, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USE THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06

; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-5

Query Match 20.6%; Score 111; DI 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 2e-104;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

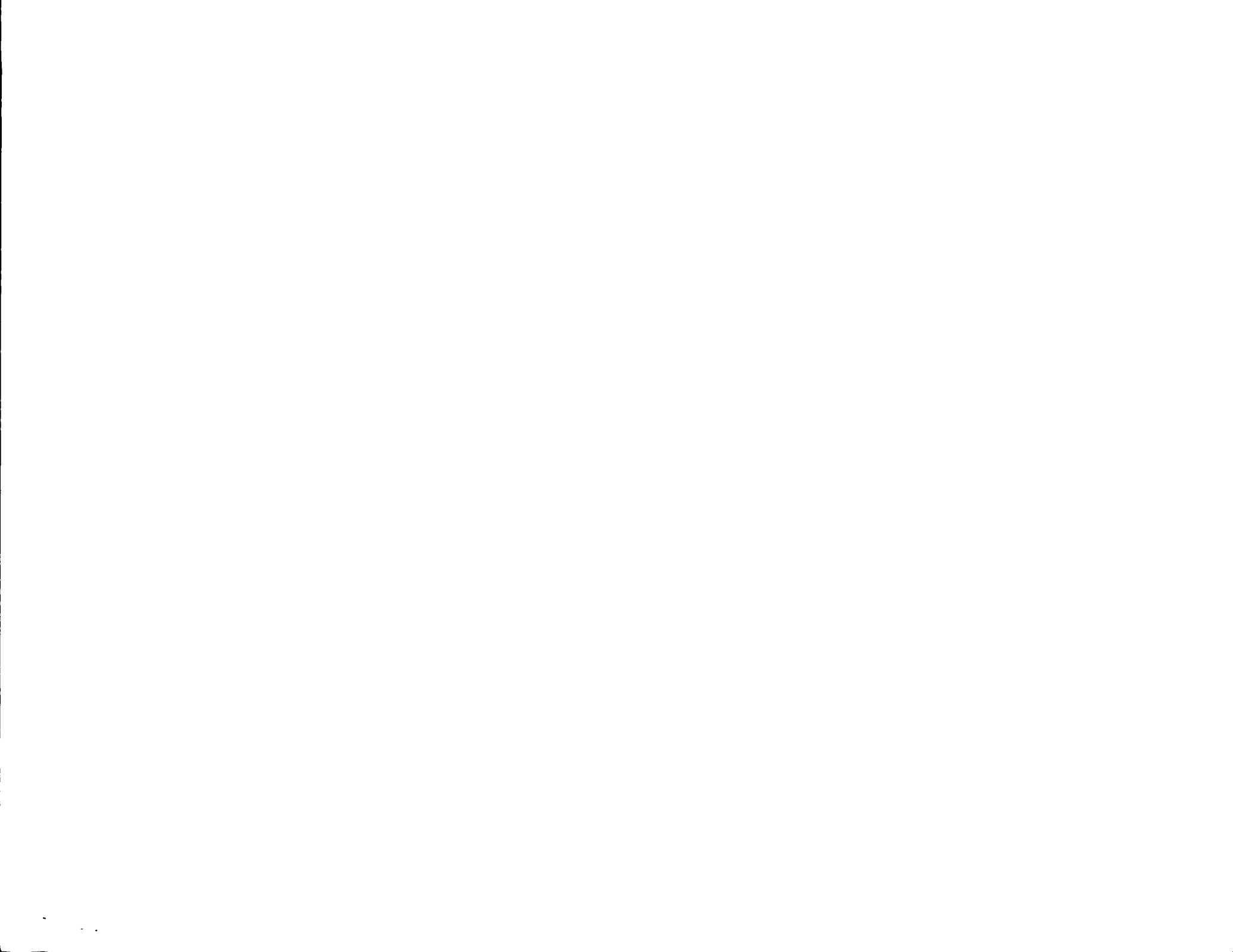
QY 301 TPIFAVIQIKKTKIQSVSSAHLCKKKYMFSLNIPVNHGQREDFVGGQLHENGSGSPET 360
Db 1 TPIFAVIQIKKTKIQSVSSAHLCKKKYMFSLNIPVNHGQREDFVGGQLHENGSGSPET 60
QY 361 SRSLPAPQDNDFLSKQAQDCYFMKLIHHCPCGNHSDSTISGSQRAAFCDHKIT 411
Db 61 SRSLPAPQDNDFLSKQAQDCYFMKLIHHCPCGNHSDSTISGSQRAAFCDHKIT 111

RESULT 12
US-09-099-041A-6
; Sequence 6, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USE THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-099-041A-6

Query Match 20.2%; Score 109; DI 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.8e-162;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 432 QPGIAQOWIQSKREIVNQMTAEACLNQSLDALLSRDLINKEDYELVSTKPTRTSKVRQLL 491
Db 1 QPGIAQOWIQSKREIVNQMTAEACLNQSLDALLSRDLINKEDYELVSTKPTRTSKVRQLL 60
QY 492 DTTDIOGEEFAKVIVQKLNKQMQCLQPYPEILVVSRSISLNLLQNKSM 540
Db 61 DTTDIOGEEFAKVIVQKLNKQMQCLQPYPEILVVSRSISLNLLQNKSM 109

RESULT 13
US-07-857-224B-25
; Sequence 25, Application US/07857224B
; Patent No. 5958784
; GENERAL INFORMATION:
; APPLICANT: Benner, Steven A.
; TITLE OF INVENTION: Predicting Folded Structures of Proteins
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Steven A. Benner
; STREET: Hadlaubstrasse 151
; CITY: Zurich
; STATE: none
; COUNTRY: Switzerland
; ZIP: (note: this is an international post code) 80602
; COMPUTER READABLE FORM:



Tue May 14 08:27:48 2002

us-09-445-223-1.ra

Page 6

MEDIUM TYPE: 3.5 inch diskette, 1.4 M storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08,701,191A
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 257
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
FEATURE: Protein Kinase: Table 8 Column 18
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-2248-25

Query Match 1.7%; Score 9; DB 2; Length 257;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADFGLS 168
DB 139 VKIADFGLS 147

RESULT 14
US-08-701-191A-0
Sequence 30, Application US/08701191A
Patent No. 542428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlesinger,
APPLICANT: and Stevan K. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 643 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08,701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227,088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-30

Query Match 1.7%; Score 9; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADFGLS 168
DB 178 VKIADFGLS 186

RESULT 15
PCT-US95-08493-2
Sequence 2, Application PC/TUS9508493
GENERAL INFORMATION:
APPLICANT: Wood, Clive
APPLICANT: Caruso, Anthony
TITLE OF INVENTION: Novel mlt Receptor Tyrosine Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08493
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: GI5234A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-08493-2

Query Match 1.7%; Score 9; DB 5; Length 530;
Best Local Similarity 100.0%; Pred. No. 0.61;

GenPore version 1.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using sw model

Run on: May 13, 2002, 19:58:15, Search time 52.18 seconds
(without alignments)
248.022 Million cell updates/sec

Title: US-09-445-223-1
Perfect score: 540
Sequence: 1 MNGPATGSAIPFIPYKELAP.....PEILVSRSPSLNLLQKSM 540

Scoring table: 0.4105
Gapop 60.0, Gapext 60.0

Searched: 231628 seqs, 2442594 residues

Word size: 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: listing first 40 summaries

Database: Issued_Patents_AA:*
1: 0002_6-prodata_1_100_5A_GOMR pep.*
2: 0002_6-prodata_1_100_5E_GOMR pep.*
3: 0002_6-prodata_1_100_6A_GOMR pep.*
4: 0002_6-prodata_1_100_6E_GOMR pep.*
5: 0002_6-prodata_1_100_6B_GOMR pep.*
6: 0002_6-prodata_1_100_6C_GOMR pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	461	85.4	478	4	US-09-069-013-4
2	461	85.4	530	4	US-09-069-013-3
3	461	85.4	531	4	US-09-069-013-1
4	461	85.4	540	1	US-09-069-013-27
5	350	66.7	540	3	US-09-019-062-1
6	260	66.7	540	4	US-09-059-01A-2
7	284	52.6	284	4	US-09-069-013-5
8	221	40.3	366	4	US-09-069-01A-4
9	195	36.1	264	4	US-09-069-013-7
10	167	30.9	167	4	US-09-069-013-8
11	111	20.6	131	4	US-09-099-061A-5
12	109	20.2	109	4	US-09-099-061A-6
13	9	1.7	257	2	US-07-857-2048-25
14	9	1.7	304	2	US-08-341-01A-30
15	9	1.7	530	5	PCT-US95-08-93-2
16	9	1.7	633	4	US-08-357-0640-43
17	9	1.7	846	1	US-08-357-068-12
18	9	1.7	946	2	US-08-357-068-12
19	9	1.7	846	5	PCT-US95-16-35-12
20	9	1.7	860	5	PCT-US95-08-93-19
21	9	1.7	868	1	US-08-374-8-4-1
22	9	1.7	869	2	US-08-374-8-4-1
23	9	1.7	868	5	PCT-US95-08-93-21
24	9	1.7	869	1	US-08-374-8-4-16
25	9	1.7	869	2	US-08-374-8-4-16
26	9	1.7	946	5	PCT-US95-03-93-13
27	9	1.7	1064	1	US-08-357-068-2

28	9	1.7	1064	2	US-09-003-289-2	Sequence 2, Appli
29	9	1.7	1064	5	PCT-US95-16445-2	Sequence 2, Appli
30	9	1.7	1082	1	US-08-357-598-5	Sequence 5, Appli
31	9	1.7	1082	1	US-08-357-598-10	Sequence 10, Appli
32	9	1.7	1082	2	US-09-003-289-5	Sequence 5, Appli
33	9	1.7	1082	2	US-09-003-289-10	Sequence 10, Appli
34	9	1.7	1082	5	PCT-US95-16445-5	Sequence 5, Appli
35	9	1.7	1082	5	PCT-US95-16445-10	Sequence 10, Appli
36	9	1.7	1098	4	US-08-946-994-17	Sequence 17, Appli
37	9	1.7	1099	4	US-08-665-5740-16	Sequence 16, Appli
38	9	1.7	1100	1	US-08-357-598-11	Sequence 11, Appli
39	9	1.7	1100	2	US-09-003-289-11	Sequence 11, Appli
40	9	1.7	1100	5	PCT-US95-16445-11	Sequence 11, Appli
41	9	1.7	1299	5	PCT-US95-08354A-2	Sequence 2, Appli
42	8	1.5	63	3	US-09-101-146-15	Sequence 15, Appli
43	8	1.5	50	2	US-08-469-537A-8	Sequence 8, Appli
44	8	1.5	50	2	US-08-469-537A-26	Sequence 26, Appli
45	8	1.5	64	1	US-08-278-089A-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-09-069-023-4
; Sequence 4, Application US/090602A
; Patent No. 6348573
; GENERAL INFORMATION:
; APPLICANT: Nurex, Gabriel
; APPLICANT: Inohara, Naohiro
; APPLICANT: Kozaki, Takayoshi
; TITLE OF INVENTION: COMPOSITION AND METHOD FOR INDUCING APOPTOSIS
; TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
; PRIORITY REFERENCE: JP 92234
; PRIORITY REFERENCE: JP 92234
; CURRENT PRIORITY DATE: 1999-04-27
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-069-023-4

Query Match 85.4%; Score 461; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q	50	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	139
DB	18	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	77
Q	140	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	199
DB	78	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	137
Q	260	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	259
DB	139	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	197
Q	280	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	319
DB	158	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	257
Q	320	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	479
DB	258	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	417
Q	380	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	439
DB	318	PIGLDHHHLEKGVVLYKMLKNNLKKKLEFVAVWLEKLEHETALGVNVLNMT	477

27 440 TC KKKEDVNMEDATINQSLDALLSRHIMKDYELVSKIKRISKVQQLDITDIOGE 499
 10 440 TC KKKEDVNMEDATINQSLDALLSRHIMKDYELVSKIKRISKVQQLDITDIOGE 499
 27 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540
 10 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540
 27 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540
 10 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540

RESULT 2

US 09-069-023-1

Sequence 1, Application US/09069023A
 Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel
 APPLICANT: Inohara, Naohiro
 APPLICANT: Kosaki, Takayoshi
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
 TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
 FILE REFERENCE: UM-03333
 CURRENT APPLICATION NUMBER: US/09-069-023A
 CURRENT FILING DATE: 1998-04-27
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 1
 LENGTH: 531
 TYPE: PR1
 ORGANISM: Homo sapiens
 US 09-069-023-1

Query Match: 85.4%; Score: 461; DB: 4; Length: 531;
 Best Local Similarity: 100.0%; Prod. No. 0;
 Matches: 461; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

27 80 PILGICNEPEFLGIVIEYMPNOSLNELLHRKTEYDDVAVWLFPRILHETALGVNYLHMI 139
 10 80 PILGICNEPEFLGIVIEYMPNOSLNELLHRKTEYDDVAVWLFPRILHETALGVNYLHMI 139
 27 140 PELLHHLKTONLEENEFHVKTADFGLSKWKMMSLSQSSKSKSAPEGGIITYMPENYE 199
 10 140 PELLHHLKTONLEENEFHVKTADFGLSKWKMMSLSQSSKSKSAPEGGIITYMPENYE 199
 27 200 PQCKSPASTKHPIYSYAVITWEVLSREGEEDFVENPLQIMYSVSDGHRPVINEESLPYD 259
 10 200 PQCKSPASTKHPIYSYAVITWEVLSREGEEDFVENPLQIMYSVSDGHRPVINEESLPYD 259
 27 250 PHFARMILLSWAGNDSRSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLST 319
 10 250 PHFARMILLSWAGNDSRSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLST 319
 27 320 ATEHTEKEMELSLNIFVNHGQERSSSQDENSSQDENSSQDENSSQDENSSQDENSSQ 379
 10 320 ATEHTEKEMELSLNIFVNHGQERSSSQDENSSQDENSSQDENSSQDENSSQDENSSQ 379
 27 380 CYMKIHHFTGNBSWISLDRSQRAALDIKTEHNSATINPSTAGNSERLQPTAQW 439
 10 380 CYMKIHHFTGNBSWISLDRSQRAALDIKTEHNSATINPSTAGNSERLQPTAQW 439
 27 440 TC KKKEDVNMEDATINQSLDALLSRHIMKDYELVSKIKRISKVQQLDITDIOGE 499
 10 440 TC KKKEDVNMEDATINQSLDALLSRHIMKDYELVSKIKRISKVQQLDITDIOGE 499
 27 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540
 10 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540

RESULT 3

US 09-069-023-1

Sequence 1, Application US/09069023A
 Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel
 APPLICANT: Inohara, Naohiro
 APPLICANT: Kosaki, Takayoshi
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
 TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
 FILE REFERENCE: UM-03333
 CURRENT APPLICATION NUMBER: US/09-069-023A
 CURRENT FILING DATE: 1998-04-27
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 27
 LENGTH: 540
 TYPE: PR1
 ORGANISM: Homo sapiens
 US 09-069-023-1

APPLICANT: Nunez, Gabriel
 APPLICANT: Inohara, Naohiro
 APPLICANT: Kosaki, Takayoshi
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
 TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
 FILE REFERENCE: UM-03333
 CURRENT APPLICATION NUMBER: US/09-069-023A
 CURRENT FILING DATE: 1998-04-27
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 1
 LENGTH: 531
 TYPE: PR1
 ORGANISM: Homo sapiens
 US 09-069-023-1

Query Match: 85.4%; Score: 461; DB: 4; Length: 531;
 Best Local Similarity: 100.0%; Prod. No. 0;
 Matches: 461; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

27 80 PILGICNEPEFLGIVIEYMPNOSLNELLHRKTEYDDVAVWLFPRILHETALGVNYLHMI 139
 10 80 PILGICNEPEFLGIVIEYMPNOSLNELLHRKTEYDDVAVWLFPRILHETALGVNYLHMI 139
 27 140 PELLHHLKTONLEENEFHVKTADFGLSKWKMMSLSQSSKSKSAPEGGIITYMPENYE 199
 10 140 PELLHHLKTONLEENEFHVKTADFGLSKWKMMSLSQSSKSKSAPEGGIITYMPENYE 199
 27 200 PQCKSPASTKHPIYSYAVITWEVLSREGEEDFVENPLQIMYSVSDGHRPVINEESLPYD 259
 10 200 PQCKSPASTKHPIYSYAVITWEVLSREGEEDFVENPLQIMYSVSDGHRPVINEESLPYD 259
 27 250 PHFARMILLSWAGNDSRSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLST 319
 10 250 PHFARMILLSWAGNDSRSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLSTLST 319
 27 320 ATEHTEKEMELSLNIFVNHGQERSSSQDENSSQDENSSQDENSSQDENSSQDENSSQ 379
 10 320 ATEHTEKEMELSLNIFVNHGQERSSSQDENSSQDENSSQDENSSQDENSSQDENSSQ 379
 27 380 CYMKIHHFTGNBSWISLDRSQRAALDIKTEHNSATINPSTAGNSERLQPTAQW 439
 10 380 CYMKIHHFTGNBSWISLDRSQRAALDIKTEHNSATINPSTAGNSERLQPTAQW 439
 27 440 TC KKKEDVNMEDATINQSLDALLSRHIMKDYELVSKIKRISKVQQLDITDIOGE 499
 10 440 TC KKKEDVNMEDATINQSLDALLSRHIMKDYELVSKIKRISKVQQLDITDIOGE 499
 27 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540
 10 500 EEAKVIVQLKKNKQMLQPYPELVVSSSPSNILLQNEFM 540

RESULT 4

US 09-069-023-27

Sequence 27, Application US/09069023A
 Patent No. 6348573

GENERAL INFORMATION:

APPLICANT: Nunez, Gabriel
 APPLICANT: Inohara, Naohiro
 APPLICANT: Kosaki, Takayoshi
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
 TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
 FILE REFERENCE: UM-03333
 CURRENT APPLICATION NUMBER: US/09-069-023A
 CURRENT FILING DATE: 1998-04-27
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO: 27
 LENGTH: 540
 TYPE: PR1
 ORGANISM: Homo sapiens
 US 09-069-023-1

```

: GENERAL INFORMATION:
: APPLICANT: Bertin, John
: TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
: TITLE OF INVENTION: DOMAIN POLYPEPTIDES
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110 2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/2010,610
: FILING DATE: 06-FEB-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Meiklejohn, Ph.D., Anita L.
: REGISTRATION NUMBER: 35,283
: REFERENCE CHECKED NUMBER: 073447069001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070

```

Query Match: 66.7%; Score: 360; DB: 4; Length: 540;
 Best Local Similarity: 99.8%; Pred. No.: 0;
 Matches: 460; Conservative: 6; Mismatches: 1; Indels: 0; Gaps: 0;

```

QY 80 PDLGTCNEPEFGIVTEYMPNGSINELLBRKTEYFDVAVWIRKELTRILHIALGVNYLHNM1 149
DB 80 PDLGTCNEPEFGIVTEYMPNGSINELLBRKTEYFDVAVWIRKELTRILHIALGVNYLHNM1 149

QY 140 PDLHHHDKTQNTLLDNETHVKIADFCLEKWMMSLSQSGESAPAGSTIYMPENYE 199
DB 140 PDLHHHDKTQNTLLDNETHVKIADFCLEKWMMSLSQSGESAPAGSTIYMPENYE 199

QY 200 PDLKRRASKKHIDYSSAVITWVLSEKKQFTEIVNGLIMYSVSGHHHPVINERSLPYDI 259
DB 200 PDLKRRASKKHIDYSSAVITWVLSEKKQFTEIVNGLIMYSVSGHHHPVINERSLPYDI 259

QY 260 PHARMISLIESGWAQNPDERSFLKCLIELEPVLRTEFEI 300
DB 260 PHARMISLIESGWAQNPDERSFLKCLIELEPVLRTEFEI 300

QY 360 ATHTCDKEEMELSLFVNBHIDELSLGQHLNPLTTEKSPAGQMBSEKKA 379
DB 360 ATHTCDKEEMELSLFVNBHIDELSLGQHLNPLTTEKSPAGQMBSEKKA 379

QY 380 CYMKLHHPTNNHWSSTLSGSKAAAFDHKTEPSSALINPLSTAGNSERLQPTTAQW 439
DB 380 CYMKLHHPTNNHWSSTLSGSKAAAFDHKTEPSSALINPLSTAGNSERLQPTTAQW 439

QY 440 TCKKPEIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQGE 499
DB 440 TCKKPEIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQGE 499

QY 540 ELAKVIVRLKINKQMNLQPYHETIVVSESTINLLNKM 549
DB 540 ELAKVIVRLKINKQMNLQPYHETIVVSESTINLLNKM 549

```

```

RESULT 7
US-09-069-023-5
: Sequence 5, Application US/09069023A
: Patent No. 6,448,573
: GENERAL INFORMATION:
: APPLICANT: Nunez, Gabriel
: APPLICANT: Inohara, Naohiro
: APPLICANT: Koseki, Takayoshi
: TITLE OF INVENTION: NOVEL MOLECULES OF THE CASPASE RELATED
: TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
: FILE REFERENCE: UM-03333
: CURRENT APPLICATION NUMBER: US/09069023A
: CURRENT FILING DATE: 1998-04-27
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO:
: LENGTH: 391
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-069-023-5

Query Match 54.9%, Score 284, DB 4, Length 284,
Best Local Similarity 100.0%, Prod. No. 4.5e-279,
Matches 284, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 340 YLAKVIVRLKINKQMNLQPYHETIVVSESTINLLNKM 349
DB 340 YLAKVIVRLKINKQMNLQPYHETIVVSESTINLLNKM 349

QY 360 ATHTCDKEEMELSLFVNBHIDELSLGQHLNPLTTEKSPAGQMBSEKKA 379
DB 360 ATHTCDKEEMELSLFVNBHIDELSLGQHLNPLTTEKSPAGQMBSEKKA 379

QY 380 CYMKLHHPTNNHWSSTLSGSKAAAFDHKTEPSSALINPLSTAGNSERLQPTTAQW 439
DB 380 CYMKLHHPTNNHWSSTLSGSKAAAFDHKTEPSSALINPLSTAGNSERLQPTTAQW 439

QY 440 TCKKPEIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQGE 499
DB 440 TCKKPEIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQGE 499

```

```

DB 160 QWTQSKREIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQ 249

QY 160 QWTQSKREIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQ 249
DB 240 QWTQSKREIVNQMTRAVLNQSLDALSPDLMKKEIYELVSTKPTFSKVRLQIDDTTQ 249

```

```

RESULT 8
US-09-099-041A-4
: Sequence 4, Application US/09099041A
: Patent No. 6,440,576
: GENERAL INFORMATION:
: APPLICANT: Berlin, John
: TITLE OF INVENTION: NOVEL MOLECULES OF THE CASPASE RELATED
: TITLE OF INVENTION: 18-TEIN FAMILY AND USES THEREOF
: FILE REFERENCE: 07134_076001
: CURRENT APPLICATION NUMBER: US/09099041A
: CURRENT FILING DATE: 1998-06-17
: US PATENT APPLICATION NUMBER: 09/016,512
: PRIOR FILING DATE: 1998-02-06
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 4
: LENGTH: 400
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-099-041A-4

```

```

Query Match 40.9%, Score 221, DB 4, Length 400,
Best Local Similarity 100.0%, Prod. No. 2.5e-215,
Matches 221, Conservative 0, Mismatches 0, Indels 0, Gaps 0

```

```

QY 80 PDLGTCNEPEFGIVTEYMPNGSINELLBRKTEYFDVAVWIRKELTRILHIALGVNYLHNM1 149
DB 80 PDLGTCNEPEFGIVTEYMPNGSINELLBRKTEYFDVAVWIRKELTRILHIALGVNYLHNM1 149

QY 140 PDLHHHDKTQNTLLDNETHVKIADFCLEKWMMSLSQSGESAPAGSTIYMPENYE 199
DB 140 PDLHHHDKTQNTLLDNETHVKIADFCLEKWMMSLSQSGESAPAGSTIYMPENYE 199

QY 200 PDLKRRASKKHIDYSSAVITWVLSEKKQFTEIVNGLIMYSVSGHHHPVINERSLPYDI 259
DB 200 PDLKRRASKKHIDYSSAVITWVLSEKKQFTEIVNGLIMYSVSGHHHPVINERSLPYDI 259

QY 260 PHARMISLIESGWAQNPDERSFLKCLIELEPVLRTEFEI 300
DB 260 PHARMISLIESGWAQNPDERSFLKCLIELEPVLRTEFEI 300

```

```

RESULT 9
US-09-069-023-7
: Sequence 7, Application US/09069023A
: Patent No. 6,448,573
: GENERAL INFORMATION:
: APPLICANT: Nunez, Gabriel
: APPLICANT: Inohara, Naohiro
: APPLICANT: Koseki, Takayoshi
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
: TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
: FILE REFERENCE: UM-03333
: CURRENT APPLICATION NUMBER: US/09069023A
: CURRENT FILING DATE: 1998-04-27
: NUMBER OF SEQ ID NOS: 38
: SOFTWARE: Patent In Ver. 2.0
: SEQ ID NO 7
: LENGTH: 264
: TYPE: PR1
: ORGANISM: Homo sapiens
US-09-069-023-7

```



```

: RESULT 14
: US-07-857 224B-25
: Sequence 25, Application 95/07857224B
: Patent No. 5958784
: GENERAL INFORMATION:
: APPLICANT: Benner, Steven A.
: TITLE OF INVENTION: Predicting Folded Structures of Proteins
: NUMBER OF SEQUENCES: 114
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Steven A. Benner
: STREET: Hadlaubstrasse 151
: CITY: Zurich
: STATE: none
: COUNTRY: Switzerland
: EMail (enter this in an International post code): CH8052
: COMPUTER-READABLE FORM:

```

1 MEDIUM TYPE: 3.5 inch diskette, 1.44 MB storage
 2 COMPUTER: Apple Macintosh
 3 OPERATING SYSTEM: Macintosh 7.0
 4 SOFTWARE: Microsoft Word
 5 CURRENT APPLICATION DATA:
 6 APPLICATION NUMBER: US/09/445-223
 7 FILING DATE: 06/21/99
 8 CLASSIFICATION: 4.0
 9 PRIOR APPLICATION DATA: none
 10 TELECOMMUNICATION INFORMATION:
 11 TELEPHONE: (International) 41 1 942 2930
 12 TELEFAX: (International) 41 1 942 2127
 13 TELEX: none
 14 INFORMATION FOR SEQ ID NO: 25:
 15 SEQUENCE CHARACTERISTICS:
 16 LENGTH: 257
 17 TYPE: amino acid
 18 TOPOLOGY: linear
 19 MOLECULE TYPE:
 20 DESCRIPTION: protein
 21 ORGANISM: Streptomyces coelicolor
 22 FEATURE: Protein Kinase; Table 8 Column 28
 23 PUBLICATION INFORMATION:
 24 AUTHOR:
 25 AUTHOR: Banks, R. K.
 26 AUTHOR: Smith, A. M.
 27 AUTHOR: Hunter, L.
 28 TITLE: The protein kinase family
 29 JOURNAL: Science
 30 VOLUME: 241
 31 PAGE: 42-52
 32 DATE: 1988
 33 SEQ ID NO: 2, 10, 25

Query Match 1.7% Score 9; DB 2; Length 257;
 Best Local Similarity 100.0%; Pred. No. 0.52;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

15 160 VKIADPGLS 168
 16 |
 17 169 VKIADPGLS 177

RESULT 14
 1 US-09-445-223-1.1A 30
 2 Sequence 1: Application US/09/445-223
 3 Patent No. 55474A
 4 GENERAL INFORMATION:
 5 APPLICANT: Moosa Mohammadi, Joseph Schlossinger,
 6 APPLICANT: and Steven R. Ballard
 7 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
 8 TITLE OF INVENTION: OF NON INHIBIN RECEPTOR TYROSINE KINASE
 9 NUMBER OF SEQUENCES: 41
 10 CORRESPONDENCE ADDRESS:
 11 ADDRESSEE: Lyon & Lyon
 12 STREET: 600 West Fifth Street
 13 STREET: Suite 4700
 14 CITY: Los Angeles
 15 STATE: California
 16 COUNTRY: U.S.A.
 17 ZIP: 90071-2066
 18 COMPUTER READABLE FORM:
 19 MEDIUM TYPE: 3.5 inch diskette, 1.44 MB
 20 MEDIUM TYPE: storage
 21 COMPUTER: IBM PC compatible
 22 OPERATING SYSTEM: IBM PC, DOS 5.0
 23 SOFTWARE: FastSeq for Windows 2.0
 24 CURRENT APPLICATION DATA:
 25 APPLICATION NUMBER: US/09/445-223-1.1A
 26 FILING DATE: August 21, 1999
 27 CLASSIFICATION: 5.0

1 PRIOR APPLICATION DATA:
 2 APPLICATION NUMBER:
 3 FILING DATE:
 4 ATTORNEY/AGENT INFORMATION:
 5 NAME: Warburg, Richard J.
 6 REGISTRATION NUMBER: 32,327
 7 REFERENCE/DOCKET NUMBER: 67,088
 8 TELECOMMUNICATION INFORMATION:
 9 TELEPHONE: (214) 489-1600
 10 TELEFAX: (214) 955-0440
 11 TELEX: 67 0510
 12 INFORMATION FOR SEQ ID NO: 30:
 13 SEQUENCE CHARACTERISTICS:
 14 LENGTH: 304 amino acids
 15 TYPE: amino acid
 16 STRANDEDNESS: single
 17 TOPOLOGY: linear
 18 MOLECULE TYPE: protein
 19 US-09-701-191A 30

Query Match 1.7% Score 9; DB 2; Length 304;
 Best Local Similarity 100.0%; Pred. No. 0.61;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADPGLS 168
 |
 DB 178 VKIADPGLS 186

RESULT 15
 1 US-09-445-223-1.1A 30
 2 Sequence 2: Application US/09/445-223
 3 GENERAL INFORMATION:
 4 APPLICANT: Wood, Olive
 5 APPLICANT: Caruso, Anthony
 6 TITLE OF INVENTION: Novel mtk Receptor Tyrosine Kinases
 7 NUMBER OF SEQUENCES: 21
 8 CORRESPONDENCE ADDRESS:
 9 ADDRESSEE: LEGAL AFFAIRS
 10 STREET: 87 CambridgePark Drive
 11 CITY: Cambridge
 12 STATE: MA
 13 COUNTRY: USA
 14 ZIP: 02140
 15 COMPUTER READABLE FORM:
 16 MEDIUM TYPE: Floppy disk
 17 COMPUTER: IBM PC compatible
 18 OPERATING SYSTEM: PC DOS/MS-DOS
 19 SOFTWARE: Patent In Release #1.0, Version #1.25
 20 CURRENT APPLICATION DATA:
 21 APPLICATION NUMBER: US/09/445-223-1.1A
 22 FILING DATE:
 23 CLASSIFICATION:
 24 ATTORNEY/AGENT INFORMATION:
 25 NAME: Brown, Scott A.
 26 REGISTRATION NUMBER: 42,724
 27 REFERENCE/DOCKET NUMBER: 61524A
 28 TELECOMMUNICATION INFORMATION:
 29 TELEPHONE: (617) 498-8224
 30 TELEFAX: (617) 876-5851
 31 INFORMATION FOR SEQ ID NO: 2:
 32 SEQUENCE CHARACTERISTICS:
 33 LENGTH: 530 amino acids
 34 TYPE: amino acid
 35 TOPOLOGY: linear
 36 MOLECULE TYPE: protein
 37 US-09-445-223-1.1A 30

Query Match 1.7% Score 9; DB 5; Length 530;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Tue May 14 08:27:48 2002

us-09-445-223-1.rai

Page 7

QY 160 VKIADFGLS 168
 |||||
DB 400 VKIADFGLS 408

Search completed: May 13, 2002, 21:33:41
Job time: 5726 sec

377
401
401
401

Query Match 36.1%; Score 195; DB 4; Length 264;
Best Local Similarity 100.0%; Pred. No. 4.3;-189;
Matches 195; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVA VPLRFRILHEIALGVNYLHNMT 139
DB 70 PILGICNEPEFLGIVTEYMPNGSLNELLHRKTEYPDVA VPI PFRILHEIALGVNYLHNMT 129
QY 140 PPLLHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPEYNE 199
DB 130 PPLLHDLKTQNILLDNEFHVKIADFGLSKWRMMSLSQSRSSKSAPEGGTIIYMPPEYNE 189
QY 200 PGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLCIMYSVSGQHRPVIINEESLPYDI 259
DB 190 PGQKSRASIKHDIYSYAVITWEVLSRKQPFEDVTNPLCIMYSVSGQHRPVIINEESLPYDI 249
QY 260 PHRARMISLIESGWA 274
DB 250 PHRARMISLIESGWA 264

PLT 10
09-069-023-6
Sequence 6, Application US/09069023A
Patent No. 6348573
GENERAL INFORMATION:
APPLICANT: Nunez, Gabriel
APPLICANT: Inohara, Naohiro
APPLICANT: Koseki, Takeyoshi
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING APOPTOSIS
TITLE OF INVENTION: SIGNALING PATHWAY INHIBITORS AND ACTIVATORS
FILE REFERENCE: UM-03333
CURRENT APPLICATION NUMBER: US/09/069,023A
CURRENT FILING DATE: 1998-04-27
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 6
LENGTH: 167
TYPE: PRT
ORGANISM: Homo sapiens
US-09-069-023-6

Query Match 30.9%; Score 167; DB 4; Length 167;
Best Local Similarity 100.0%; Pred. No. 5.3e-161;
Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 374 SRKAQDCYFMKLHHCPCGNHSDSTISGSQRAAFCDHKITPCSSAIIINPLSTAGNSERLQP 433
DB 1 SRKAQDCYFMKLHHCPCGNHSDSTISGSQRAAFCDHKITPCSSAIIINPLSTAGNSERLQP 60
QY 434 GIAQOWIQSKREDIVNQMTAEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVROLLE 493
DB 61 GIAQOWIQSKREDIVNQMTAEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVROLLE 120
QY 494 IDIQGEFFAKVIVOKLKDNKQMGLOPYPEIIVVSRSEINLNLQKSM 540
DB 121 IDIQGEFFAKVIVOKLKDNKQMGLOPYPEIIVVSRSEINLNLQKSM 167

RESULT 11
US-09-099-041A-5
Sequence 5, Application US/09099041A
Patent No. 6340576
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USE THEREOF
FILE REFERENCE: 07334-076001
CURRENT APPLICATION NUMBER: US/09/099,041A
CURRENT FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 09/019,942
PRIOR FILING DATE: 1998-02-06

NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 5
LENGTH: 131
TYPE: PRT
ORGANISM: Homo sapiens
US-09-099-041A-5

Query Match 20.6%; Score 111; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 2e-104;
Matches 111; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 TEIFAVIQIKKTKIQSVSAIHCCKKKMELSLNIPVNHGPOEESCGGSLHENSQSPET 360
DB 1 TEIFAVIQIKKTKIQSVSAIHCCKKKMELSLNIPVNHGPOEESCGGSLHENSQSPET 60
QY 361 SRSLPAPQDNDFLSKAQDCYFMKLHHCPCGNHSDSTISGSQRAAFCDHKT 411
DB 61 SRSLPAPQDNDFLSKAQDCYFMKLHHCPCGNHSDSTISGSQRAAFCDHKT 111

RESULT 12
US-09-099-041A-6
Sequence 6, Application US/09099041A
Patent No. 6340576
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USE THEREOF
FILE REFERENCE: 07334-076001
CURRENT APPLICATION NUMBER: US/09/099,041A
CURRENT FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 37
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 6
LENGTH: 109
TYPE: PRT
ORGANISM: Homo sapiens
US-09-099-041A-6

Query Match 20.2%; Score 109; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 1.8e-102;
Matches 109; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 432 QPGIAQOWIQSKREDIVNQMTAEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVROLL 491
DB 1 QPGIAQOWIQSKREDIVNQMTAEACLNQSLDALLSRDLIMKEDYELVSTKPTRTSKVROLL 60
QY 492 IDTDIQGEFFAKVIVOKLKDNKQMGLOPYPEIIVVSRSEINLNLQKSM 540
DB 61 IDTDIQGEFFAKVIVOKLKDNKQMGLOPYPEIIVVSRSEINLNLQKSM 109

RESULT 13
US-07-857-224B-25
Sequence 25, Application US/07857224B
Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 15
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8002
COMPUTER READABLE FORM:



Tue May 14 08:27:48 2002

us-09-445-223-1.ra

Page 6

MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: 08/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2437
TELEFAX: (International) 41 1 252 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 257
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: Saccharomyces cerevisiae
FEATURE: Protein kinase; Table 8 Column 2;
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1986
US-07-857-224B-25

Query Match 1.7%; Score 9; DB 1; Length 257;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADFGLS 168
|||||
139 VKIADFGLS 147

RESULT 14
US-08-701-191A-30
Sequence 30, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlessinger,
APPLICANT: and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 555-5440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 304 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-30

Query Match 1.7%; Score 9; DB 2; Length 304;
Best Local Similarity 100.0%; Pred. No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 VKIADFGLS 168
|||||
Db 178 VKIADFGLS 186

RESULT 15
PCT-US95-08493-2
Sequence 2, Application PC/TUS9508493
GENERAL INFORMATION:
APPLICANT: Wood, Clive
APPLICANT: Caruso, Anthony
TITLE OF INVENTION: Novel mlk Receptor Tyrosine Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEGAL AFFAIRS
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0; Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08493
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: G15234A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-08493-2

Query Match 1.7%; Score 9; DB 5; Length 530;
Best Local Similarity 100.0%; Pred. No. 0.61;



RESULTS
 AA092795
 Title: AAK02795: Homo sapiens: Fc gamma R1A.
 XX
 Accession: AAK02795;
 XX
 Date: 07 MAY 1998 (10:16:00)
 XX
 Description: Homo: B1 protein.
 DE
 XX
 FW: B1 protein, intracellular domain,
 FW: cytosolic tail path., intracellular
 XX
 OS: Homo sapiens.
 XX
 PN: W0985507.A2.
 XX
 PD: 10 DEC 1998.
 XX
 UF: 01 JUN 1998; 98w01100255.
 XX
 PR: 11 SEP 1997; 9711-0121746.
 PR: 05-JUN-1997; 9711-0121011.
 PR: 03 JUN 1997; 9711-0121199.
 XX
 PA: (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI: Boldin M, Mariani N, Wallach D;
 XX
 GK: Wilk, 1998 (970266/06).
 LE: N LTRB. AAK02795.
 XX
 FT: Bx-B1 protein regulates cell death
 PT: derivatives, DNA and antibodies,

E1 Information for treating AIDS, Cancer
 XX
 PS Title: 1. F1. 3A. 5ppp: English.
 XX
 XX
 C1 This invention describes the isolation of a novel human B1 protein which
 C2 can interact with intracellular mediators or modulators of inflammation,
 C3 cell death and/or cell survival pathways, directly or indirectly. Cells
 C4 can be modulated or mediated in inflammation, cell death or cell survival
 C5 pathways or another intracellular signalling activity using B1.
 C6 Conditions such as AIDS and cancer can be treated using B1. Antibodies,
 C7 oligonucleotides and ribozymes can also be used to regulate the above
 C8 pathways.
 XX
 S1 Sequence: 540 AA:

Query Match: 100.0%; Score: 540; DB: 20; Length: 540;
 Best Local Similarity: 100.0%; Pred. No.: 0;
 Matches: 540; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0.

27 1 MNDEALVALPIPYHKIADLPYSEKASSTVSSARHAEWQVAVVHEHHHHHLLISSEK 60
 DB 1 MNDEALVALPIPYHKIADLPYSEKASSTVSSARHAEWQVAVVHEHHHHHLLISSEK 60
 97 91 QAVRFAADLHKARESYIEPTLADLNNEFEEIVVYMENGEHLLHKKIEYPTVAVWL 120
 DB 91 QAVRFAADLHKARESYIEPTLADLNNEFEEIVVYMENGEHLLHKKIEYPTVAVWL 120
 98 91 QAVRFAADLHKARESYIEPTLADLNNEFEEIVVYMENGEHLLHKKIEYPTVAVWL 120
 DB 91 QAVRFAADLHKARESYIEPTLADLNNEFEEIVVYMENGEHLLHKKIEYPTVAVWL 120
 99 121 LKKELHETALGVNYLHNMIPPELHDEKIQNLEENEFVYLADEGLSKWMMSSQSSES 180
 DB 121 LKKELHETALGVNYLHNMIPPELHDEKIQNLEENEFVYLADEGLSKWMMSSQSSES 180
 100 161 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 240
 DB 161 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 240
 101 241 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 300
 DB 241 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 300
 102 301 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 360
 DB 301 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 360
 103 361 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 420
 DB 361 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 420
 104 421 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 480
 DB 421 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 480
 105 481 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 540
 DB 481 LKSAPEEELIYMPPENYEGRKASIKRILYSYAVITWEVLSEKQPTDVINILQIMY 540

B03M1 1.2
 AAY594.1
 C1 AAY594.1 (human); Protein: 478 AA
 XX
 A1 AAY594.1
 XX
 C1 MAR 1.0 (first entry)
 XX
 C1 Human B1 protein sequence residues 54-541.
 XX
 KW B1 protein (human); B1-like interacting (B1AR) kinase; apoptosis regulator; ARK;
 KW caspase-8; caspase-9; CD95 signalling; apoptosis signalling pathway;
 KW CD95; ARK; B1; DRK; 1; diagnosis; cell growth; apoptosis; dysregulation;
 KW per therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
 KW apoptosis; ischaemic injury; toxin induced liver disease.
 KW

PS Homo sapiens.
 XX
 PS W09955134-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 27-APR-1999; 99WO/0509183.
 XX
 PR 27-APR-1998; 98US/0069024.
 XX
 PA (DNM1) DNIV MICHIGAN.
 XX
 PI Nunez G, Inohara N, Koseki I;
 XX
 DR WP1: 2000-072163/06.
 XX
 PI Compositions for identifying apoptosis signalling pathway inhibitors
 PI useful for treating diseases
 XX
 PS Claim 6; Page: 5; 5ppp: English.
 XX
 C1 This sequence is a fragment of the human B1 (B1-like interacting (B1AR)
 C2 kinase) protein of the invention. B1AR acts as a positive regulator
 C3 of apoptosis, potentiating apoptosis induced by caspase 8 and caspase 9
 C4 during CD95 signalling. The invention provides methods for identifying
 C5 apoptosis signalling pathway inhibitors and activators, and methods and
 C6 compositions for screening compounds which will modulate the interactions
 C7 of the various compositions identified: ARK, B1, and the CIDE family of
 C8 activators (CIDE A, CIDE B and CIDE C). B1 is useful in screening
 C9 assays for agents, useful in the diagnosis, prognosis or treatment of
 C10 disease associated with excess cell growth and dysregulation of
 C11 apoptosis. Complexes containing B1 and B1AR can be used in drug
 C12 screening assays to identify inhibitor molecules blocking CD95-mediated
 C13 apoptosis. Overexpression of ARK in an in vitro cell system can be used
 C14 to identify inhibitors of the enzymatic activity of caspase 8.
 C15 Identification of ARK-like inhibitory compounds may be useful for gene
 C16 therapy treatment of disease with increased cell death in muscle tissue
 C17 and cardiac disorders. Therapeutic compositions of CIDEs can be used to
 C18 treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
 C19 ischaemic injury, and toxin induced liver disease. AntiB1 antibodies
 C20 can be used as reagents for the preparation or affinity chromatography
 C21 media, and for diagnostically measuring B1 levels. A specific inhibitor
 C22 of an essential step in the biochemistry of apoptosis is needed. B1
 C23 interaction with intracellular factors such as B1AR and FADD appears to
 C24 be essential for apoptosis. Inhibitors of B1 binding to intracellular
 C25 apoptosis factors are potential drug candidates.
 C26 Note: This sequence was created using information given in the
 C27 specification.
 XX
 S1 Sequence: 478 AA:

Query Match: 85.4%; Score: 461; DB: 21; Length: 478;
 Best Local Similarity: 100.0%; Pred. No.: 0;
 Matches: 461; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0.

97 80 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 120
 DB 80 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 120
 98 181 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 240
 DB 181 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 240
 99 241 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 300
 DB 241 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 300
 100 301 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 360
 DB 301 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 360
 101 361 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 420
 DB 361 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 420
 102 421 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 480
 DB 421 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 480
 103 481 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 540
 DB 481 PLGLGNEPEPELIVTEYHFRGLHLLHKKIEYPTVAVWLQAVLHETALGVNYLHNM 540

```

Db      258  ahhleddkfkveisihpwhap-pes-svssqphes-qspetsrlgagqdndtlrkaqd 317
QY      380  CYEMKLIHHCENHNSWDSTLSQRAAPCHHTTPTCSAATINFI STAGNSERLQPGIAQOW 439
Db      319  /tshllhepnh-wdsrlsagpntshkttgssallqsltagusertlgslagqw 377
QY      440  IQSKREELIVNMTIFACINQSIDALLSERIMKREYHIVSTKPTPTSEVPEQLDITTEIQGE 499
Db      478  lqskrediveqpteanlqpsidallspdlmkedylcstskptitakvrgliddtdiqge 437
QY      500  RFKRVIVQKLEIKKQMCIQYFEILVVSRSPLAILQNKSM 540
Db      438  etakvinkkldakqmglyppellvvsrsplailqnksm 478

```

```

RESULT 3
AA59404
ID   AAY59404 standard; Protein: 541 AA
XX
AC   AAY59404;
XX
DI   21-MAY-2000 (first entry)
XX
DE   Human RICK protein sequence.
XX
KW   RICK; human; RIP-like; interaction; CIAPF kinase; apoptosis; regulation; AEC;
KW   caspase-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
KW   CIDE-A; CIDE-B; DPPF-1; diagnosis; cell growth; apoptosis; dysregulation;
KW   gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
KW   aplastic anaemia; ischaemic injury; alcohol-induced liver disease
XX
OS   Homo sapiens.
XX
PR   W09551:1 A2.
XX
PD   04-NOV-1999.
XX
PF   27-APR-1999 4490 DBS9193
XX
PR   27-APR-1999 9808 DBS9021.
XX
PA   (UNM) UNIV MICHIGAN
XX
PI   Nishida G, Inokubo N, Kozaki T.
XX
DR   WPI: 2000:72164/06.
DR   N-PSDB: AAY48762.
XX
PT   Compositions for identifying apoptosis signalling pathway inhibitors
XX   useful for treating diseases -
XX
PS   Claim 1: Fig 7a: 93pp; English.
XX

```

```

CC   This sequence is the human RICK (RIP-like interacting CIAPF kinase)
CC   protein of the invention. The RICK protein acts as a positive regulator
CC   of apoptosis, potentiating apoptosis induced by caspase 8 and caspase-10
CC   during CD95 signalling. The invention provides methods for identifying
CC   apoptosis signalling pathway inhibitors and activators, and methods and
CC   compositions for screening compounds which will modulate the interactions
CC   of the various compositions identified. AEC, RICK, and the CIDE family of
CC   activators (CIDE-A, CIDE-B and DPPF-1). RICK is useful in screening
CC   assays for agents, useful in the diagnosis, prognosis or treatment of
CC   disease associated with excess cell growth and dysregulation of
CC   apoptosis. Complexes containing RICK and CIAPF can be used in drug
CC   screening assays to identify inhibitor molecules blocking CD95 mediated
CC   apoptosis. Overexpression of AEC in an in vitro cell system can be used
CC   to identify inhibitors of the enzymatic activity of caspase-8.
CC   Identification of AEC-like inhibitory compounds may be useful for gene
CC   therapy treatment of disease with increased cell death in muscle tissue
CC   and cardiac disorders. Therapeutic compositions of CIDEs can be used to
CC   treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
CC   ischaemic injury, and alcohol-induced liver disease. Anti-RICK antibodies

```

```

CC   can be used as reagents for the preparation or affinity chromatography
CC   media, and for diagnostically measuring RICK levels. A specific inhibitor
CC   of an essential step in the biochemistry of apoptosis is needed. RICK
CC   interaction with intracellular factors such as CIAPF and FADD appears to
CC   be essential for apoptosis. Inhibitors of RICK binding to intracellular
CC   apoptosis factors are potential drug candidates.
XX
SQ   Sequence 531 AA;

```

```

Query Match      85.4%; Score 461; Db 21; Length 531;
Best Local Similarity 100.0%; Prod. No. 0;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      80  PILGIENEPETIAITGYPPAPSLNELHKKTEYPDRAWPLFRLVLSLIRNYLHMT 139
DB      71  pilgienepetiaitgyppapslnelhkkteyprawplfrlvlslirnylhmt 130
QY      140  PHLHHHKLKFNILLNHFHVEIAFSLFKEWEMHPIQYFKEZCAHPSNCTIYHNPENYE 199
DB      131  pphlhhklkfpilldnfhvkiadrlakwrmmslqsyekskapeatliymppenye 190
QY      200  PAVEKPAKAKNLYSYAVLWVYLSRRGHEFIVTNHPTMYVSVQCHRPVINEESLEYDI 259
DB      191  pavgkpasaknllysyavlwvylsrrghfivtnhptmyvsvqchrvineesleydi 250
QY      260  ILSAEMISLITLWAGNIIISLITLITLITLITLITLITLITLITLITLITLITLITLIT 319
DB      251  ghsamslitlswagndiisliitlililililililililililililililililil 310
QY      320  ARLGKRRKMLLRLIVRGLIQLSRLQVHHSLSLRLKSLALQWFLSEEAQI 379
DB      311  ahlldkkmellrlivrgllqlslrlqvhhsllslrlksslalqwfllseeaqi 370
QY      380  CYEMKLIHHCENHNSWDSTLSQRAAPCHHTTPTCSAATINFI STAGNSERLQPGIAQOW 439
DB      371  cyemklihhepnhswdstlsqraapchhthtptcsaatinfi stagnserlqpgiaqow 430
QY      440  IQSKREELIVNMTIFACINQSIDALLSERIMKREYHIVSTKPTPTSEVPEQLDITTEIQGE 499
DB      431  lqskrediveqpteanlqpsidallspdlmkedylcstskptitakvrgliddtdiqge 490
QY      500  RFKRVIVQKLEIKKQMCIQYFEILVVSRSPLAILQNKSM 540
DB      491  rfkrvivqklekqmcqyfeilvvsrsplailqnksm 540

```

```

RESULT 4
AA568774
ID   AAY568774 standard; Protein: 541 AA.
XX
AC   AAY568774;
XX
DI   16 MAY-2000 (first entry)
XX
DE   Amino acid sequence of a human phosphorylation effector PHSP-6.
XX
KW   Human; phosphorylation effector; PHSP; proliferative disorders;
KW   immune disorder; neuronal disorder.
XX
OS   Homo sapiens.
XX
FH   Key Location/Qualifiers
FT   Region 18..287
FT   Modified-site 23 /note "protein kinase family signature sequence"
FT   Modified-site 34 /note "potential phosphorylation site"
FT   Modified-site 34 /note "potential phosphorylation site"
FT   Modified-site 58 /note "potential phosphorylation site"
FT   Modified-site 100 /note "potential phosphorylation site"
FT   Modified-site 100 /note "potential glycosylation site"

```

FI Modified site: 192
 FI Znote: "potential phosphorylation site"
 FI Modified site: 196
 FI Znote: "potential phosphorylation site"
 FI Modified site: 198
 FI Znote: "potential phosphorylation site"
 FI Modified site: 207
 FI Znote: "potential phosphorylation site"
 FI Modified site: 224
 FI Znote: "potential phosphorylation site"
 FI Modified site: 267
 FI Znote: "potential phosphorylation site"
 FI Modified site: 296
 FI Znote: "potential phosphorylation site"
 FI Modified site: 301
 FI Znote: "potential phosphorylation site"
 FI Modified site: 360
 FI Znote: "potential phosphorylation site"
 FI Modified site: 373
 FI Znote: "potential phosphorylation site"
 FI Modified site: 374
 FI Znote: "potential phosphorylation site"
 FI Modified site: 391
 FI Znote: "potential phosphorylation site"
 FI Modified site: 401
 FI Znote: "potential phosphorylation site"
 FI Modified site: 428
 FI Znote: "potential phosphorylation site"
 FI Modified site: 442
 FI Znote: "potential phosphorylation site"
 FI Modified site: 457
 FI Znote: "potential phosphorylation site"
 FI Modified site: 478
 FI Znote: "potential phosphorylation site"
 FI Modified site: 478
 FI Znote: "potential phosphorylation site"
 FI Modified site: 484
 FI Znote: "potential phosphorylation site"
 FI Modified site: 537
 FI Znote: "potential phosphorylation site"
 XX
 GN WO2000/5450-A1
 XX
 DE 08 FEB 2000
 XX
 CL 18 JUL 2000; 56WO/051713
 XX
 CR 18 JUL 2000; 560US 0124404
 CR 14 SEP 2000; 560US 0152014
 CR 14 OCT 2000; 560US 0174402
 CR 03 NOV 2000; 560US 0106009
 CR 10 NOV 2000; 560US 0109004
 CR 27 DEC 2000; 560US 0114796
 CR 17 JAN 2000; 560US 0174402
 CR 12 JAN 2000; 560US 0229005
 XX
 CA (INCYT, INC.): PHARM, INC.
 XX
 CI Brillman M, Gail P, Tang YF, Corley NC, Guedler KJ, Baughn MR,
 PI Patterson C, Bandman O, Au Young J, Gersone GA, Yue H, Azimzai Y,
 CI Brady R, De AM, Shih JL
 XX
 CR WO 99/200189, 25/7/99
 CR N 99080; AAL4644
 XX
 CI New human phosphorylation effectors useful for the diagnosis, treatment
 and prevention of proliferative, immune and neuronal disorders
 XX
 DE Claim 1: Comp 94-95; 142pp; English
 XX
 CR AAY676/97 and AAY687/97 represent human phosphorylation effectors
 (PHSP). Isolated PHSP1-PHSP4 (the protein sequence for PHSP28 is not
 given in the specification). The sequences were isolated from cDNA
 libraries prepared from various human tissues. The PHSP proteins are
 useful for the diagnosis, treatment and prevention of proliferative

CC disorders, immune disorders and neuronal disorders. The PHSP proteins
 CC form pharmaceutical compositions which useful for treating or preventing
 CC disorders associated with decreased PHSP expression/activity. PHSP
 CC antagonists are useful for treating or preventing disorders associated
 CC with increased PHSP expression/activity.
 XX
 SQ Sequence: 540 AA;

Query Match: 85.48%; Score 461; DB 21; Length 540;
 Best Local Similarity: 100.00%; Pred. No: 0;
 Matches: 461; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

QY 80 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 139
 DB 80 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 139
 QY 140 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 199
 DB 140 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 199
 QY 200 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 259
 DB 200 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 259
 QY 260 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 319
 DB 260 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 319
 QY 320 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 379
 DB 320 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 379
 QY 380 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 439
 DB 380 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 439
 QY 440 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 499
 DB 440 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 499
 QY 500 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 540
 DB 500 PHLGPNHRTHLAVTYEYMNQNNHLLRKRTHYDVAWHLRKRTHLHLAGVNYHNM 540

RESULT: 5
 AAB4 5570
 ID AAB4 5570 standard; Protein: 544 AA;
 XX
 AC AAB4 5570;
 XX
 DT 08 FEB 2001 (first entry)
 XX
 DE Human cancer associated protein sequence SEQ ID NO:1015.
 XX
 KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nontropic;
 KW vasotropic; antipsoriasis; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; hematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host diseases; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening;
 XX
 OS Homo Sapiens.
 XX
 PN WO2000/5450-A1
 XX
 PD 21-SEP 2000.
 XX

CC This invention describes the isolation of novel human caspase
CC recruitment domain (CARD) 3 and CARD 4 polynucleotide, cDNAs and proteins and a
CC partial murine (Akk4) protein and genes, the genes and proteins of
CC the invention are involved in the regulation of caspase activation.
CC The caspase recruitment domain (CARD) polynucleotides, polypeptides,
CC homodimers and antibodies can be used in screening assays, detection
CC assays, predictive medicine and therapeutic and prophylactic methods of
CC treatment. The methods may be used to diagnose and treat patients which
CC are suffering from a disorder associated with abnormal level or rate of
CC apoptotic cell death, abnormal activity of the Fas/CD95 receptor
CC complex, abnormal activity of the TNF receptor complex, or abnormal
CC activity of a caspase. Disclosed that can be treated in lab can or
CC (particularly follicular lymphoma, carcinomas associated with mutations

in pituitary hormone-dependent tumours), autoimmune disorders (e.g., systemic lupus erythematosus, immune-mediated glomerulonephritis), viral infections, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, retinitis pigmentosa, spinal muscular dystrophy, cerebellar degeneration, anemias, myelodysplastic syndrome, myocardial infarction, and stroke. CARD-3 protein interacts with other cellular proteins, and so can be used for regulation of cellular proliferation and differentiation and cell survival. The CARD proteins may also be used to for screen drugs or compounds which modulate their activity. The CARD-4 gene can express a long transcript that encodes CARD-4L, a short transcript that encodes CARD-4S or two CARD-4 splice variants, CARD-4Y and CARD-4Z. This sequence represents the human CARD-3 protein described in the method of the invention.

XX
SQ Sequence 540 AA

Query Match 66.7%; Score 360; DB 22; Length 540;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 80 FIIIGTGNIEPEPLGIVIEYKPNQSLNELLHRKTHYDVAVWLKRFLLHMLALGVNYLHMMT 139
DB 80 piliitgnepetlgizteympnaslnellhrkteydvavwplrtllhmlalgvnylhmmt 139
QY 140 PPIILHDLKIQNLILNIEHVKIADPGISKEWMMMSISQSSKSKSAPGGTIIYMPENYK 199
DB 140 ppiilhdldkqnlildnethkaidtqskwmmsslsqssksksapggdtiymppenyk 199
QY 200 PPKKSFASKKHDIYSYAVITWEVLSRKQPEEDVINPILQIMYSVSGHRIIVINEESLPYDI 259
DB 200 pppkksfaskkhdiiysyavitwevlsrkqptedtvinpilhqimysvsgchriivineeslpydi 259
QY 260 PHFARMISLIESQWAGNIDHRSFLEKTELEIVLPTFEETFEFAVIGLKKTKLQSVSS 319
DB 260 phfarmisliesqwaqndrpsflekteleivlptfeetfeefaviglkktklqsvss 319
QY 320 AILHLDKKEMMLNITFVNRHILPQVITVHLHNSLIPHTDSLEFAPDNLNHLKKAQ 379
DB 320 ailhldkkemmlnitfvnrhildpvtvhlhnsllphtdslefapdnlnhlkkaq 379
QY 380 YILMKIADCHPNRSWISLIESQFAARQHKILPNSALINPLSLATNSERIQPGIAGQW 439
DB 380 yilmkiadchpnrswidsliesqfaarqhkilpnssalinplslatnseriqpgiagqw 439
QY 440 LSKKRLIVNKHILATINLHSAVILFRLMKKHTHVLWTEKPKTSKVRQLLHITLQK 499
DB 440 lskkrlivnkhilatinlhhsavilfrlmmkhtthvlwtepktskvrqlhltlqk 499
QY 500 LSKAVIVLKRLNKKMILPYPHILVWVSRPLNIDQKSM 540
DB 500 lskavivlkrlnkkmilpyphilvwvsrplnidqksm 540

KEYWORDS
AAK02001

DB AAK02001 (19-04-01) 540 AA

XX
AT AAK02001

XX
GI 15 AAF0001 (19-04-01) 540

XX
DE Human CARD-3 protein

XX
KW CARD-3; apoptosis; recruitment; disease; human; cancer; infection;
XX
KW autoimmune diseases; neurological diseases; haematological diseases;
XX
KW immune diseases; inflammation; and tumours; and isleptin;
XX
KW immunomodulators; and inflammation; apoptosis; diagnosis;
XX
KW disease

XX
CT Human Sequence

XX
GB Key word: CARD-3 protein

E1 Domain 1..400
E1 /note: "kinase domain"
E1 Domain 401..431
E1 /note: "linker domain"
E1 Domain 432..540
E1 /note: "CARD"

XX
FN W0200100826-A2

XX
FD 04-JAN-2001

XX
FF 28 JUN 2000; 2000WO 0S17691

XX
FR 28 JUN 1999; 99DS 0340620

XX
JA (MILL.) MILLENNIUM PHARM INC.

XX
E1 Berlin JJ

XX
DE WPI; 2001 06197 (207)

XX
DE N-PSDB; AAF0001

XX
E1 Isolated intracellular proteins predicted to be involved in regulating
E1 caspase activation are used for diagnosis and treatment of e.g. cancer,
E1 viral infections, autoimmune diseases, neurological diseases and
E1 haematological disorders -

XX
PS Claim 9; Fig 2; 200pp; English

XX
CC The present sequence is that of human caspase recruitment domain-3
CC (CARD-3), an intracellular protein predicted to be involved in
CC regulating caspase activation. The sequence is predicted from an
CC isolated cDNA clone (see AAF0001). Methods of diagnosing and
CC treating patients suffering from a disorder associated with an
CC abnormal level or rate of apoptotic cell death, abnormal activity
CC of the Fas/AP0-1 receptor complex, abnormal activity of the tumour
CC necrosis factor receptor complex or abnormal activity of a caspase
CC involve administering a compound that modulates the expression or
CC activity of CARD-3, CARD-4, CARD-5 or CARD-6 e.g. using gene
CC therapy methods. Such disorders include cancer, viral infection,
CC autoimmune disorders, neurological diseases, haematological
CC disorders, inflammatory disorders and immune disorders. CARD-3,
CC 4, -5 and -6 proteins can be used to regulate cell proliferation,
CC cell survival and cell growth. They can also be used to screen
CC drugs or compounds that modulate their activity or expression and
CC to treat disorders associated with insufficient or excessive
CC production of CARD-3, -4, -5 or -6 protein, or production of an
CC aberrant protein.

XX
SQ Sequence 540 AA

Query Match 66.7%; Score 360; DB 22; Length 540;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 460; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 80 FIIIGTGNIEPEPLGIVIEYKPNQSLNELLHRKTHYDVAVWLKRFLLHMLALGVNYLHMMT 139
DB 80 piliitgnepetlgizteympnaslnellhrkteydvavwplrtllhmlalgvnylhmmt 139
QY 140 PPIILHDLKIQNLILNIEHVKIADPGISKEWMMMSISQSSKSKSAPGGTIIYMPENYK 199
DB 140 ppiilhdldkqnlildnethkaidtqskwmmsslsqssksksapggdtiymppenyk 199
QY 200 PPKKSFASKKHDIYSYAVITWEVLSRKQPEEDVINPILQIMYSVSGHRIIVINEESLPYDI 259
DB 200 pppkksfaskkhdiiysyavitwevlsrkqptedtvinpilhqimysvsgchriivineeslpydi 259
QY 260 PHFARMISLIESQWAGNIDHRSFLEKTELEIVLPTFEETFEFAVIGLKKTKLQSVSS 319
DB 260 phfarmisliesqwaqndrpsflekteleivlptfeetfeefaviglkktklqsvss 319
QY 320 AILHLDKKEMMLNITFVNRHILPQVITVHLHNSLIPHTDSLEFAPDNLNHLKKAQ 379

```

Db 320 ahhedkkkmelsinipvohgpggesegsqahensgspetssipapqndfllrkaqd 379
Qy 380 CYFMKLHHCPGNHSDSTISCSQRAAFCDHKHTPCSSAIINPLSTAGNSERLQPGIAQQW 439
Db 380 cyfmklhhcpghswdstisgsqraafcdhktipcssaiinplstagnserlqpgiaqqw 439
Qy 440 IQSKKEDI VNMTEACLNQSLDALLSRDLIMKEDYELVSTEPIRTSEKVPQLLITDIOGE 499
Db 440 iqsksredivngateacnlqslaldllsrdlmkedyelvtstkprrtskvrqlldttidige 499
Qy 500 PPAKVIVQKIKDNKQMGLOPYPRILVVSRSPLNLIKONKSM 540
Db 500 efaqvivqkikdnkmglopyprilvvsrpslnlikonksm 540

```

RESULT 8
AAY59406

ID AAY59406 standard: Protein: 284 AA.

XX
AC AAY59406;

XX
DT 21-MAR-2000 (first entry)

XX
DE Human RICK protein sequence residues 248-531.

XX
KW RICK: human; RIP-like interacting CLARP kinase; apoptosis regulator; APC;
caspace-8; caspase-10; CD95 signalling; apoptosis signalling pathway;
CIDP-A; CIDP-B; DREP-1; diagnosis; cell growth; apoptosis dysregulation;
gene therapy; cardiac disorder; cancer; neurodegenerative disorder; AIDS;
aplastic anaemia; ischaemic injury; toxin-induced liver disease

XX
OS Homo sapiens.

XX
PN W09955134-A2.

XX
PD 04-NOV-1999.

XX
PF 27-APR-1999; 99WO-0509183.

XX
PR 27-APR-1998; 98US-0069023.

XX
PA (UNM1) UNIV MICHIGAN.

XX
PI Nunez G, Inohara N, Koseki T;

XX
DR WPI: 2000-072163/06.

XX
PT Compositions for identifying apoptosis signalling pathway inhibitors
useful for treating diseases -

XX
PS Claim 5: Page -; 93pp; English.

XX
CC This sequence is a fragment of the human RICK (RIP-like interacting CLARP
kinase) protein of the invention. RICK acts as a positive regulator
of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10
during CD95 signalling. The invention provides methods for identifying
apoptosis signalling pathway inhibitors and activators, and methods and
compositions for screening compounds which will modulate the interactions
of the various compositions identified: ARC, RICK, and the CIDP family of
activators (CIDP-A, CIDP-B and DREP-1). RICK is useful in screening
assays for agents, useful in the diagnosis, prognosis or treatment of
disease associated with excess cell growth and dysregulation of
apoptosis. Complexes containing RICK and CLARP can be used in drug
screening assays to identify inhibitor molecules blocking CD95-mediated
apoptosis. Overexpression of ARC in an in vitro cell system can be used
to identify inhibitors of the enzymatic activity of caspase-8
identification of ARC-like inhibitory compounds may be useful for gene
therapy treatment of disease with increased cell death in muscle tissue
and cardiac disorders. Therapeutic compositions of CIDPs can be used to
treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia,
ischaemic injury, and toxin-induced liver disease. AntiRICK antibodies
can be used as reagents for the preparation or affinity chromatography

CC media, and for diagnostically measuring RICK levels. A specific inhibitor
of an essential step in the biochemistry of apoptosis is needed. RICK
interaction with intracellular factors such as CLARP and FADD appears to
be essential for apoptosis, inhibitors of RICK binding to intracellular
apoptosis factors are potential drug candidates.
CC Note: This sequence was created using information given in the
CC specification.

XX
SO Sequence 284 AA;

Query Match 52.6% Score 284; DB 21; Length 284;
Best local Similarity 100.0%; Prod. No. 1.3e-284;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 257 YDIPHRAPMINSIESGWAQNPDERPSFLKCLIELEPVLTFTFETLEAVIQLYKTKLOS 316
Db 1 ydiphrapmnsiesgwaqnpderpsflkclielepvtftfettleaviqlykktkllqs 60
Qy 317 VASAELDALEFHELELNIVYMHGAPTPGSSGRIHNGCSPTSPSPAPQDNDELSEK 376
Db 61 vssaihledkkkmelsinipvohgpggesegsqahensgspetssipapqndfllrkaqd 120
Qy 377 AQDYFMKLHHCPGNHSDSTISGSQRAAFCDHKHTPCSSAIINPLSTAGNSERLQPGIA 426
Db 121 aqdyfmklhhcpghswdstisgsqraafcdhktipcssaiinplstagnserlqpgia 180
Qy 427 QWLVSEKPIVNMTEACLNQSLDALLSRDLIMKEDYELVSTEPIRTSEKVPQLLITDIO 496
Db 181 qwlvsekpivngateacnlqslaldllsrdlmkedyelvtstkprrtskvrqlldttid 240
Qy 497 PPAKVIVQKIKDNKQMGLOPYPRILVVSRSPLNLIKONKSM 540
Db 241 ppaqvivqkikdnkmglopyprilvvsrpslnlikonksm 284

```

RESULT 9

AAM93621

ID AAM93621 standard: Protein: 540 AA.

XX
AC AAM93621;

XX
DT 06-NOV-2001 (first entry)

XX
DE Human polypeptide, SEQ ID NO: 3454.

XX
KW human; full length cDNA; cDNA synthesis; oligo-capping.

XX
OS Homo sapiens.

XX
PN EP1140094-A2.

XX
PD 05-SEP-2001.

XX
PF 07-001-2000; 2000EP-0114089.

XX
PR 08-JUL-1999; 99JP-0194486.

XX
PR 11-JAN-2000; 2000JP-0118774.

XX
PR 02-MAY-2000; 2000JP-0183765.

XX
PA (HELIX) HELIX RES INST.

XX
PI Oda T, Nishikawa T, Inada T, Hayashi K, Ishii S, Kawai Y,
Wakamatsu A, Sugiama T, Nagai K, Kojima M, Ohtsuki S, Koda H;

XX
DR WPI: 2001-524255/58.

XX
DR N-PSDB: AAK94554.

XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
use in genetic manipulation -

XX
PS Claim 8: SEQ ID NO 3454; 1380pp + sequence listing; English.

XX

PD 04-Nov-1999.
XX
PF 27-APR-1999; 99WO/0509183.
XX
PR 27-APR-1998; 98US/0069023.
XX
PA (UNRE) UNIV MICHIGAN.
XX
PI Nurey G., Inohara N., Koseki T.
XX
DR WP1; 2000/072163/86.
XX
PT Compositions for identifying apoptosis signalling pathway inhibitors
PI useful for treating diseases
XX
PS Claim 6; Page : 93pp; English.

This sequence is a fragment of the human RICK (RIP-like interacting cIARP kinase) protein of the invention. RICK acts as a positive regulator of apoptosis, potentiating apoptosis induced by caspase-8 and caspase-10 during CD95 signalling. The invention provides methods for identifying apoptosis signalling pathway inhibitors and activators, and methods and compositions for screening compounds which will modulate the interactions of the various compositions identified: ABC, RICK, and the CIDE family of activators (CIDE-A, CIDE-B and CIDE-1). RICK is useful in screening assays for agents, useful in the diagnosis, prognosis or treatment of disease associated with excess cell growth and dysregulation of apoptosis. Complexes containing RICK and CIDE can be used in drug screening assays to identify inhibitor molecules blocking CD95-mediated apoptosis. Overexpression of ABC in an in vitro cell system can be used to identify inhibitors of the enzymatic activity of caspase-8. Identification of ABC-like inhibitory compounds may be useful for gene therapy treatment of disease with increased cell death in muscle tissue and cardiac disorders. Therapeutic compositions of CIDEs can be used to treat e.g. cancer, AIDS, neurodegenerative disorders, aplastic anaemia, ischaemic injury, and toxin-induced liver disease. Anti-RICK antibodies can be used as reagents for the preparation of affinity chromatography media, and for diagnostically measuring RICK levels. A specific inhibitor of an essential step in the biochemistry of apoptosis is needed. RICK interaction with intracellular factors such as CIDE and FADD appears to be essential for apoptosis, inhibitors of RICK binding to intracellular apoptosis factors are potential drug candidates.
 Note: This sequence was created using information given in the specification.

Query Match 30.9%, Score 167, DB sz: Length 167;
 Best Local Similarity 100.0%, Pred. No. 2.3e-163;
 Matches 167; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible][illegible]

```

RESULT 11
AAB58938
ID AAB58938 standard; Protein; 154 AA
XX
AC AAB58938;
XX
DI 27 MAR 2001 (first entry)
XX

```


DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 646.

XX

KW Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
KW neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
KW antidiabetic; antiinflammatory; antileuc; vulacrar; anticonvulsant;
KW antibacterial; antifungal; antiparasitic; cardiant; immune disorder;
KW Addison's disease; allergy; autoimmune hemolytic anaemia;
KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
KW cardiovascular disorder; wound healing; neurological disease.

XX

OS Homo sapiens.

XX

PN W0200055173-11

XX

PD 21-SEP-2000.

XX

FF 08-MAR-2000; 2000WO-0205881.

XX

FR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Rosen CA, Ribon SM.

XX

DR WPI; 2000-411515/58

DR

N-PSDB; AAF21841.

XX

PT New human breast and ovarian cancer associated gene sequences and the
PT polypeptides encoded by these genes, useful in the prevention,
PT treatment and diagnosis of cancer, immune disorders, cardiovascular
PT disorders and neurological diseases -

XX

PS Claim 11; Page 1086; 1299pp; English.

XX

CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
CC proteins AAB29711 - AAB59109. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59110 which are used in the
CC isolation and characterisation of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic; immunosuppressive;
CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antileuc; vulacrar; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiant activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases

XX

SO Sequence 153 AA:

Query Match 29.3%; Score 126, DB D1, Length 153,
Best Local Similarity 100.0%; Prod. No. 3.4e-121;
Matches 126, Conservative 0, Mismatches 0, Indels 0, Gaps 0.

Oy 97 YMPNCSI NPI IRETFYDVAWHI RPR IIRTAI CANYI IIMTFFLI RHILKPTENI IEN 157

Db 1 ympnscilhrkxteydvawp lrrilheial gnylihmtpplihadlktgnilds 60

Oy 157 RPRVPIATPISLKNWMMISLSSRSKSAIPGRII YMHGNYHIDPSFASIKHHCYSYA 216

Db 61 rprvpiatpislknwmmislsrsksaipgr iymhgnyhidpsfaskhhdhysya 120

Oy 217 VITWEV 222

Db 121 vitwev 126

RESULT 12

AAM25501

ID AAM25501 standard, Protein, 142 AA.

XX

AC AAM25501;

XX

DT 16-OCT-2001 (first entry)

XX

DE Human protein sequence SEQ ID No:1016.

XX

KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiagregant; haemostatic; vulnerary; antileuc; osteopathia; eczema;
KW dermatological; antiallergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; neotropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antiaphyllactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopenia, osteoporosis, severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.

XX

OS Homo sapiens.

XX

PN W0200153455 A2.

XX

PD 26-JUL-2001.

XX

FF 22-OCT-2000; 2000WO US35017.

XX

FR 23-SEP-1999; 99US 0171275.

FR 21-JAN-2000; 2000US-0498725.

FF 25-APR-2000; 2000US 0502417.

XX

FA (HSL) BYSEQ IRL.

XX

PI Tang YL, Liu C, Erimanac RT;

XX

DR WPI; 2001-457603/49.

DR

N-PSDB; AAH99442.

XX

PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -

XX

PS Claim 20; Page 269; 1217pp; English.

XX

CC AAM99195 to AAM99901 encode the human proteins given in AAM25225 to
CC AAM25462. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiagregant; haemostatic; vulnerary;
CC antileuc; osteopathia; dermatological; antiallergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; neotropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.

XX
02 Sequence: 14; AA:

Query Match: 1.00%; Score 10; DB Zz; Length 142;
Best Local Similarity: 100.00%; Posit. Neg. 0.0006;
Matches: 0; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

27 116-DL-TGN1140-155
11-11111-1
05 154-distep11-155

RESULTS: 13

AA017205

10 AA017205 (standard); Protein: 167 AA;

XX

A1 AA017205;

XX

07 NOV 2000 (first entry)

XX

06 Novel signal transduction pathway protein; Seq ID 800;

XX

KW Neuroprotection; cytostatic; dermatological; immunosuppressive; tumour;
anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;
immune system disorder; rheumatoid arthritis; inflammatory condition;
organ transplant rejection; infection; hepatitis C; blood disorder;
sickle cell anemia; hyperproliferative disorder; Gaucher's disease;
neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
chromosomal abnormality; Down syndrome; lactacidia; renal disorder;
cardiovascular respiratory; wound healing; endocrine; Addison's disease;
reproductive system; gastrointestinal; liver disorder; AIDS;
immunodeficiency syndrome;

XX

05 Homo Sapiens;

XX

06 W00011734 A1;

XX

01 AUG 2001

XX

07 JAN 2001; 2001W01050112;

XX

01 JAN 2000; 200005 0176065;

04 FEB 2000; 200005 0180628;

24 FEB 2000; 200005 0184664;

02 MAR 2000; 200005 0189459;

16 MAR 2000; 200005 0189874;

17 MAR 2000; 200005 0190076;

18 APR 2000; 200005 0198123;

19 MAY 2000; 200005 0205515;

07 JUN 2000; 200005 0209467;

28 JUN 2000; 200005 0214806;

09 JUN 2000; 200005 0215135;

07 JUL 2000; 200005 0216647;

07 JUL 2000; 200005 0216880;

11 JUL 2000; 200005 0217487;

13 JUL 2000; 200005 0217496;

11 JUL 2000; 200005 0218290;

26 JUL 2000; 200005 0220963;

26 JUL 2000; 200005 0220964;

14 AUG 2000; 200005 0224518;

14 AUG 2000; 200005 0224519;

14 AUG 2000; 200005 0225213;

14 AUG 2000; 200005 0225214;

14 AUG 2000; 200005 0225266;

14 AUG 2000; 200005 0225267;

14 AUG 2000; 200005 0225268;

14 AUG 2000; 200005 0225270;

14 AUG 2000; 200005 0225447;

14 AUG 2000; 200005 0225757;

14 AUG 2000; 200005 0225758;

14 AUG 2000; 200005 0225759;

14 AUG 2000; 200005 0226279;

PR 22-AUG-2000; 200005-0226681;
PR 22-AUG-2000; 200005-0226868;
PR 22-AUG-2000; 200005-0227182;
PR 23-AUG-2000; 200005-0227009;
PR 30-AUG-2000; 200005-0228924;
PR 01-SEP-2000; 200005-0229287;
PR 01-SEP-2000; 200005-0229343;
PR 01-SEP-2000; 200005-0229344;
PR 01-SEP-2000; 200005-0229345;
PR 05-SEP-2000; 200005-0229509;
PR 05-SEP-2000; 200005-0229513;
PR 06-SEP-2000; 200005-0230437;
PR 06-SEP-2000; 200005-0230438;
PR 08-SEP-2000; 200005-0231242;
PR 08-SEP-2000; 200005-0231243;
PR 08-SEP-2000; 200005-0231244;
PR 08-SEP-2000; 200005-0231413;
PR 08-SEP-2000; 200005-0231414;
PR 08-SEP-2000; 200005-0232080;
PR 08-SEP-2000; 200005-0232081;
PR 12-SEP-2000; 200005-0231968;
PR 14-SEP-2000; 200005-0232497;
PR 14-SEP-2000; 200005-0232498;
PR 14-SEP-2000; 200005-0232499;
PR 14-SEP-2000; 200005-0232400;
PR 14-SEP-2000; 200005-0232401;
PR 14-SEP-2000; 200005-0233063;
PR 14-SEP-2000; 200005-0233064;
PR 14-SEP-2000; 200005-0233065;
PR 21-SEP-2000; 200005-0234223;
PR 21-SEP-2000; 200005-0234274;
PR 25-SEP-2000; 200005-0234997;
PR 25-SEP-2000; 200005-0234998;
PR 26-SEP-2000; 200005-0235484;
PR 27-SEP-2000; 200005-0235834;
PR 27-SEP-2000; 200005-0235836;
PR 29-SEP-2000; 200005-0236327;
PR 29-SEP-2000; 200005-0236367;
PR 29-SEP-2000; 200005-0236368;
PR 29-SEP-2000; 200005-0236369;
PR 29-SEP-2000; 200005-0236470;
PR 02-OCT-2000; 200005-0236802;
PR 02-OCT-2000; 200005-0237047;
PR 02-OCT-2000; 200005-0237038;
PR 02-OCT-2000; 200005-0237039;
PR 02-OCT-2000; 200005-0237040;
PR 14-OCT-2000; 200005-0239935;
PR 14-OCT-2000; 200005-0239937;
PR 20-OCT-2000; 200005-0240960;
PR 20-OCT-2000; 200005-0241221;
PR 20-OCT-2000; 200005-0241785;
PR 20-OCT-2000; 200005-0241786;
PR 20-OCT-2000; 200005-0241787;
PR 20-OCT-2000; 200005-0241808;
PR 20-OCT-2000; 200005-0241809;
PR 20-OCT-2000; 200005-0241826;
PR 01-NOV-2000; 200005-0244617;
PR 08-NOV-2000; 200005-0246474;
PR 08-NOV-2000; 200005-0246475;
PR 08-NOV-2000; 200005-0246476;
PR 08-NOV-2000; 200005-0246477;
PR 08-NOV-2000; 200005-0246478;
PR 08-NOV-2000; 200005-0246523;
PR 08-NOV-2000; 200005-0246524;
PR 08-NOV-2000; 200005-0246525;
PR 08-NOV-2000; 200005-0246526;
PR 08-NOV-2000; 200005-0246527;
PR 08-NOV-2000; 200005-0246528;
PR 08-NOV-2000; 200005-0246532;
PR 08-NOV-2000; 200005-0246609;
PR 08-NOV-2000; 200005-0246610;
PR 08-NOV-2000; 200005-0246611;
PR 08-NOV-2000; 200005-0246613;

PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249247.
PR 17-NOV-2000; 2000US-0249248.
PR 17-NOV-2000; 2000US-0249249.
PR 17-NOV-2000; 2000US-0249250.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251090.
PR 05-DEC-2000; 2000US-0256710.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254057.
PR 05-JAN-2001; 2001US-0255678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI: 2001-165460/50.
DR N-PSDB: AAS27152.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PI prognosing disorders related to the proteins, including cancers, immune
PI disorders and neuronal disorders.
XX
PS Claim 1: SEQ ID No 800; P80pp; Enlist.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejection and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders (primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies (e.g. acquired immune deficiency
CC syndrome (AIDS)). AA017056-AA017692 represent novel signal transduction
CC pathway protein amino acid sequences of the invention.
XX

Query Match: 1.0%; Score 11; Db 22; Length 167;
Best Local Similarity: 100.0%; Pred. Io: 0.1;
Matches: 10; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

W: 146 DIKTQNIID 155
D: 4 DIKTQNIID 13

RESULT: 14
AA017241
ID AA017241 standard, Protein, 254 AA.
XX
AC AA017241:
XX
DT 07 NOV 2001 (first entry)
XX
DE Novel signal transduction pathway protein, seq ID 800.
XX
KW B-cell protective; cytostatic; dermatological; immunosuppressive; tumour;
KW anti-inflammatory; anti-HIV; antibacterial; anti-inflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX
OS Homo sapiens.
XX
FN W0200104733 A1.
XX
PD 02 AUG 2001.
XX
FI 17 JAN 2001; 2001US-050142.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0190124.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 24-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344
 PR 01-SEP-2000; 2000US-0229345
 PR 05-SEP-2000; 2000US-0229509
 PR 05-SEP-2000; 2000US-0229513
 PR 06-SEP-2000; 2000US-0230437
 PR 06-SEP-2000; 2000US-0230438
 PR 06-SEP-2000; 2000US-0231242
 PR 06-SEP-2000; 2000US-0231243
 PR 06-SEP-2000; 2000US-0231244
 PR 06-SEP-2000; 2000US-0231413
 PR 06-SEP-2000; 2000US-0231414
 PR 06-SEP-2000; 2000US-0232000
 PR 06-SEP-2000; 2000US-0232001
 PR 12-SEP-2000; 2000US-0231968
 PR 14-SEP-2000; 2000US-0232397
 PR 14-SEP-2000; 2000US-0232398
 PR 14-SEP-2000; 2000US-0232399
 PR 14-SEP-2000; 2000US-0232400
 PR 14-SEP-2000; 2000US-0232401
 PR 14-SEP-2000; 2000US-0233063
 PR 14-SEP-2000; 2000US-0233064
 PR 14-SEP-2000; 2000US-0233065
 PR 21-SEP-2000; 2000US-0234223
 PR 21-SEP-2000; 2000US-0234224
 PR 25-SEP-2000; 2000US-0234597
 PR 25-SEP-2000; 2000US-0234598
 PR 26-SEP-2000; 2000US-0235404
 PR 27-SEP-2000; 2000US-0235634
 PR 27-SEP-2000; 2000US-0235635
 PR 29-SEP-2000; 2000US-0236327
 PR 29-SEP-2000; 2000US-0236328
 PR 29-SEP-2000; 2000US-0236368
 PR 29-SEP-2000; 2000US-0236369
 PR 29-SEP-2000; 2000US-0236370
 PR 02-OCT-2000; 2000US-0236802
 PR 02-OCT-2000; 2000US-0237037
 PR 02-OCT-2000; 2000US-0237038
 PR 02-OCT-2000; 2000US-0237039
 PR 02-OCT-2000; 2000US-0237040
 PR 13-OCT-2000; 2000US-0239035
 PR 13-OCT-2000; 2000US-0239037
 PR 23-OCT-2000; 2000US-0243060
 PR 23-OCT-2000; 2000US-0241221
 PR 23-OCT-2000; 2000US-0241705
 PR 23-OCT-2000; 2000US-0241706
 PR 23-OCT-2000; 2000US-0241707
 PR 23-OCT-2000; 2000US-0241809
 PR 23-OCT-2000; 2000US-0241809
 PR 23-OCT-2000; 2000US-0241809
 PR 01-NOV-2000; 2000US-0244617
 PR 08-NOV-2000; 2000US-0246374
 PR 08-NOV-2000; 2000US-0246375
 PR 08-NOV-2000; 2000US-0246476
 PR 08-NOV-2000; 2000US-0246477
 PR 08-NOV-2000; 2000US-0246478
 PR 08-NOV-2000; 2000US-0246523
 PR 08-NOV-2000; 2000US-0246524
 PR 08-NOV-2000; 2000US-0246525
 PR 08-NOV-2000; 2000US-0246526
 PR 08-NOV-2000; 2000US-0246527
 PR 08-NOV-2000; 2000US-0246528
 PR 08-NOV-2000; 2000US-0246532
 PR 08-NOV-2000; 2000US-0246609
 PR 08-NOV-2000; 2000US-0246610
 PR 08-NOV-2000; 2000US-0246611
 PR 08-NOV-2000; 2000US-0246613
 PR 17-NOV-2000; 2000US-0249207
 PR 17-NOV-2000; 2000US-0249208
 PR 17-NOV-2000; 2000US-0249209
 PR 17-NOV-2000; 2000US-0249210
 PR 17-NOV-2000; 2000US-0249211
 PR 17-NOV-2000; 2000US-0249212
 PR 17-NOV-2000; 2000US-0249213

PR 17-NOV-2000; 2000US-0249214
 PR 17-NOV-2000; 2000US-0249215
 PR 17-NOV-2000; 2000US-0249216
 PR 17-NOV-2000; 2000US-0249217
 PR 17-NOV-2000; 2000US-0249218
 PR 17-NOV-2000; 2000US-0249244
 PR 17-NOV-2000; 2000US-0249245
 PR 17-NOV-2000; 2000US-0249264
 PR 17-NOV-2000; 2000US-0249265
 PR 17-NOV-2000; 2000US-0249297
 PR 17-NOV-2000; 2000US-0249299
 PR 17-NOV-2000; 2000US-0249300
 PR 01-DEC-2000; 2000US-0250160
 PR 01-DEC-2000; 2000US-0250391
 PR 05-DEC-2000; 2000US-0251030
 PR 05-DEC-2000; 2000US-0251988
 PR 05-DEC-2000; 2000US-0256719
 PR 06-DEC-2000; 2000US-0251475
 PR 08-DEC-2000; 2000US-0251856
 PR 08-DEC-2000; 2000US-0251868
 PR 08-DEC-2000; 2000US-0251869
 PR 08-DEC-2000; 2000US-0251989
 PR 08-DEC-2000; 2000US-0251990
 PR 11-DEC-2000; 2000US-0254097
 PR 05-JAN-2001; 2001US-0259678

XX

PA (HUMA) HUMAN GENOME B11 INI.

XX

F1 Rosen CA, Barash SC, Rubin SM;

XX

DR WPI: 2001-46546/50.

DR N PSDR: AAS27158.

XX

F1 Novel polypeptides useful for diagnosing, treating, preventing and/or
 F1 prognosing disorders related to the proteins, including cancers, immune
 PT disorders and neuronal disorders

XX

PS Claim 1: SEQ ID No 806: 880pp; English.

XX

CC The invention relates to novel isolated polypeptides (I), and
 CC polynucleotides (II), (I), (II) and the antibody to (I) are useful for
 CC diagnosing, preventing and treating diseases including immune system
 CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
 CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
 CC transplant rejections and graft versus host disease, infectious diseases
 CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
 CC other blood related disorders (sickle cell anaemia), myeloproliferative
 CC disorders, primary haematopoietic disorders, hyperproliferative
 CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
 CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
 CC disorders (e.g. glomerulonephritis), cardiovascular disorders
 CC (e.g. arrhythmia), respiratory disorders, dermatological disorders, in-
 CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
 CC Addison's disease), reproductive system disorders, gastrointestinal
 CC disorder (inflammatory disorders), liver disorders (cirrhosis),
 CC as stimulators of B cell responsiveness to pathogens, activators of
 CC T-cells, to induce higher affinity antibodies, and as a means to induce
 CC tumour proliferation in pathologies e.g. acquired immune deficiency
 CC syndrome (AIDS). AA017059-AA017683 represent novel signal transduction
 CC pathway protein, amino acid sequences of the invention.

XX

Query Match 1.9%; Score 10; DB 2; Length 254;
 Best Local Similarity 100.00; Prod. No. 0.15
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 146 DLKTONHEID 155

11111111

Db 132 DLKTPHID 141

RESULT 15

AAB85038
 ID AAB85038 standard; Protein: 330 AA
 XX
 AC AAB85038;
 XX
 DI 06-AUG-2001 (first entry)
 XX
 DE Human SER4 protein sequence.
 XX
 KW Serine/threonine Kinase-like protein; SERX; breast; testicular; cancer;
 KW cardiovascular; autoimmune disorder; cytostatic; cardiac; gene therapy
 KW immunosuppressive; antiasthmatic; antifibrinolytic; SER4.
 XX
 OS Homo sapiens.
 XX
 PN W0200136645-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 17 NOV-2000; 2000WO-US31744.
 XX
 PR 17-NOV-1999; 99US-0165986.
 PR 09-FEB-2000; 2000US-0181347.
 PR 03-APR-2000; 2000US-0194195.
 PR 05-APR-2000; 2000US-0194839.
 PR 07-APR-2000; 2000US-0195647.
 PR 13-APR-2000; 2000US-0197080.
 PR 15-SEP-2000; 2000US-0232677.
 PR 16-NOV-2000; 2000US-0165986.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Quinn KE, Spytek KA, Majumder K, Verret C, Burgess C, Fernandes E;
 PI Taupier R, Rastelli L, Herrmann JL;
 XX
 DR WPI: 2001-336006/35.
 DR N-PSDB: AAF63969.
 XX
 PT New isolated SERX polynucleotides and polypeptides related to the
 PT serine/threonine kinase family of proteins, useful for determining and
 PT treating diseases related to altered levels of the polypeptide -
 XX
 PS Claim 1; Page 21; 127pp; English.
 XX
 CC The invention relates to serine/threonine kinase-like proteins, termed
 CC SERX. The SERX polynucleotides can be used to express SERX protein (e.g.
 CC via a recombinant expression vector in a host cell in gene therapy
 CC applications), to detect SERX mRNA (e.g. in a biological sample) or a
 CC genetic lesion in a SERX gene and modulate SERX activity. The SERX
 CC polypeptides, polynucleotides, antibodies can also be used to treat or
 CC prevent a pathology associated with SERX. Disorders associated with
 CC aberrant SERX expression or activity include breast or testicular cancer,
 CC cardiovascular defect, autoimmune disorders and asthma. Disorders
 CC associated with aberrant SERX expression include, blood clotting
 CC hematopoietic and tumor related disorders. The present sequence
 CC represents the amino acid sequence of human SER4 protein.
 XX
 SQ Sequence 330 AA;

Query Match 1.9%; Score 10; DB 22; Length 330;
 Best Local Similarity 100.0%; Pred No 0.14;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 146 DIKIQNILLD 155
 |||||
 Db 128 DIKIQNILLD 137

Search completed: May 13, 2002, 21:32:38
 Job time: 8628 sec



GenCore version 1.5
Copyright (c) 1992 - 2000 Ampugen Ltd

OM nucleic - nucleic search, using sw model

Run on: May 13, 2002, 19:55:39, Seq ch time 2937.39 Seconds
(without alignments)
9640.059 Million cell updates/sec

Title: us-09-445-223-2
Perfect score: 2098
Sequence: 1 ggccattatggtatggtatggg. aaaaaaaaaaaaaaaaaa 2098

Scoring table: BLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 13736267 seqs, 6748477512 residues

Word size: 0

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 266666666

Post-processing: Listing first 45 summaries

Database: EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_in:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	692	32.2	811	10	BC170405	BC170405 60242276P1
2	656	31.0	883	10	BC393551	BC393551 602411943
3	650	31.0	883	10	BC393551	BC393551 602411943
4	598	28.5	1011	10	BM469352	BM469352 AGENCOURT
5	592	28.2	592	9	AI745575	AI745575 uc341112.1
6	554	26.4	636	10	BE551615	BE551615 7442306.1
7	551	26.3	912	10	BI257472	BI257472 602967861
8	532	25.4	615	9	AA160647	AA160647 uc349811.1
9	522	25.4	647	9	AW960501	AW960501 EST772572
10	505	24.1	879	10	BE877822	BE877822 601486392
11	500	23.9	759	10	BI268990	BI268990 602972165
12	455	21.7	606	9	AW085560	AW085560 uc349811.1
13	449	21.4	455	9	AI343850	AI343850 uc349811.1
14	448	21.4	835	10	BI267776	BI267776 603034378
15	445	21.2	828	10	BF875947	BF875947 601486423
16	421	20.2	615	9	AT801150	AT801150 uc349811.1
17	423	20.2	423	9	AA574167	AA574167 uc349811.1

c 18	422	20.1	518	9	AW593657	AW593657 uc349811.1
c 19	422	20.1	541	9	AW150819	AW150819 uc349811.1
c 20	412	19.5	844	10	BM014010	BM014010 uc349811.1
c 21	411	19.6	483	9	AI343247	AI343247 uc349811.1
c 22	404	19.4	445	9	AW820715	AW820715 RC5-ST030
c 23	394	18.8	400	9	AA825052	AA825052 uc349811.1
c 24	393	18.7	487	9	AW393452	AW393452 RC5-ST030
c 25	391	18.6	1056	10	BM468817	BM468817 AGENCOURT
c 26	387	18.4	776	10	BC574192	BC574192 602596077
c 27	386	18.4	742	10	BI225743	BI225743 602251365
c 28	381	18.2	423	10	BF062840	BF062840 uc349811.1
c 29	371	17.7	1112	10	BF125423	BF125423 601763549
c 30	370	17.6	446	9	BE002713	BE002713 QV4-BN009
c 31	355	16.9	476	9	AA913804	AA913804 uc349811.1
c 32	355	16.9	616	9	AI307810	AI307810 uc349811.1
c 33	353	16.8	445	9	AW820730	AW820730 RC5-ST030
c 34	350	16.7	350	10	BF754627	BF754627 QV3-CT055
c 35	350	16.7	422	9	AA723532	AA723532 uc349811.1
c 36	349	16.6	418	9	AI499817	AI499817 uc349811.1
c 37	348	16.6	439	9	AI380994	AI380994 uc349811.1
c 38	346	16.5	426	9	AW385642	AW385642 PM0-LI003
c 39	341	16.3	456	9	AW393431	AW393431 RC5-ST030
c 40	341	16.3	945	10	BI258379	BI258379 602029440
c 41	333	15.9	446	9	AW610171	AW610171 RC5-ST030
c 42	331	15.8	616	10	BI669380	BI669380 603295259
c 43	325	15.5	429	9	AL158366	AL158366 uc349811.1
c 44	322	15.3	777	10	BI918703	BI918703 603037520
c 45	319	15.2	493	10	C05866	C05866 uc349811.1

ALIGNMENTS

RESULT 1
BC170405
LOCUS BC170405 811 bp mRNA linear EST 06-FEB-2001
DEFINITION 60242276P1 NIH_MGC_89 Homo sapiens cDNA clone IMAGE:4426016 5', mRNA sequence.
ACCESSION BC170405
VERSION BC170405.1 GI:12677108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 811)
REFERENCE
AUTHOR NIH MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: rgs@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the MGC website at:
http://imgc.llnl.gov
Plate: LLAM10172 row: 4 column: 09
High quality sequence step: 721.
FEATURES
source Location/Qualifiers
1..811
/organism "Homo sapiens"
/db_xref "taxon:9606"
/clone "IMAGE:4426016"
/clone_lib "NIH_MGC_89"
/tissue_type "hypernephroma, cell line"
/lab_host "BMDP (phage-resistant)"
/vector "pUC19 (phage-resistant)"
/site_1 "5' end, cloned unidirectionally; oligo-dT primed."
Average insert size 1.2 kb. Library enriched for full length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

oy 1619 acctgccttaaccatgcctagatgcccttctgtccagggacttgaatcatgaadgaaggac 1678

REFID: 12
 W77574.2
 1,008 nt 6,775/74,2 nt 952 bp mRNA Linear EST 15 MAY 2001
 DEFINITION: NM_011614.1 Non-MMTV 18 kDa mammary-specific cDNA clone (MAM1805142.1),
 cDNA sequence.
 ACCESSION: W77574.2
 VERSION: W77574.2.1 GI:14068075
 KEYWORDS: none
 SOURCE: human
 ORGANISM: Homo sapiens
 Eukaryota; Metazoa; Chordata; Mamalia; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (bases 1 to 952)
 REFERENCE: NIH MGC <http://www.ncbi.nlm.nih.gov/>
 TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL: Unpublished (1999)

Db 541 GGCAGGCTTAACCGAGTCGGTAGATGCCCTCTCTGTCAGGGACTTGTATCATGAAGAGGGAC 609

Qy 1679 TATGAACCTGTATAGTACCAAGGCTACCAAGGCTCAAAATGCAATCAATACAG 1734

Db 501 TATGAACCTGTATAGTACCAAGGCTACCAAGGCTCAAAATGCAATCAATACAG 556

```

RESULT 3
BC:93551
LOCUS      BC93551             883 bp      mRNA      linear      EST 12-MAR-2001
DEFINITION      F011194 Cl. NIH_MGC_92 Homo sapiens cDNA clone IMAGE:454787.57.
               mRNA sequence.
ACCESSION      BC93551
VERSION       BC93551.1      GI:11286999
KEYWORDS      EST
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1. (bases 1 to 883)
AUTHOR       GIB MC LAY, J. (Gen. et al. 1997)
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL       Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
               Email: rga@bbs.femail.nih.gov
               Tissue Procurement: ATCC
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The L.M.A.C.U. Consortium (LLNL)
               DNA Sequencing by: TruCyte Genomics, Inc.
               Clone Distribution: MGC Clone Distribution Information can be
               found through the L.M.A.C.U. Consortium, 2010 at
               http://image.llnl.gov
               Plate: LLAM10471 row: c column: 12
               High quality sequence stop: 767.

FEATURES             Location/qualifiers
     source           1..883
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4540787"
                     /clone_lib="NIH_MGC_92"
                     /tissue_type="embryonic fibroblast cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 2.5 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."

BASE COUNT      292 a      196 c      164 g      230 t      1 others
ORIGIN

```

Query Match:	31.0%:	Score 650:	DB 10:	Length 883:
Best Local Similarity	100.0%:	Pred. No. 4e-141:		
Matches: 650:	Conservative: 0:	Mismatches: 0:	Indels: 0:	Gaps: 0:

Qy 1267 ttttaattcatatggtccacaaagagaaatcatgttgatcatttcagatccatggaataatgg 1326
Db 1 TGTAAATCATGGTCCACAACAGCAATCATGCGA CCGTCAGGTCCATGAAAAATAGTGG 60

Qy 1227 TTTTCTGGAAGTTCAGAGTTCATGGAGGTCCCTACCAACTGATTCTTATAAAGAAA 1286
|||||
Db 51 TTCTCGCAAAGCTTCAGAGTCCCTGCCAGCTCCTTAAGACAATGATTCTTATCTAGAAA 120

Qy 1387 a p p c a a g a c t g t t a t t t t a t g a a g e t g c a t e a c g t c e e t g g a a t c a c a g t t g g q a t a g 1446
|||||
Db 121 A G C T C A A G A C T G T T A T T T A T G A A G E T G C A T E A C G T C E E T G G A A T C A C A G T T G G Q A T A G 180

Qy 1117 GAAATTTCTGGATGATGAAAGGATGGATTTGATATGACAGAGACACTGCATGCTCTTTC 1506
|||||
Db 181 CACCATTTCTGGATGATGAAAGGATGGATTTGATATGACAGAGACACTGCATGCTCTTTC 240

DOI: 10.1002/anie.201705293

Db 241 AGCAATAATAAAATCACTCTCAACTGCAGGAAACTCAAAAGCTCTGCAGCTGGTATAGC 300

[illegible]

DB 401 C CAGCAGTGGATCTAGAGCAAAAGCGAAGACATGGTGAAACAAATGATAGAAGCCTGCCT 360

Qy 1627 Taaaccagt cagct aagat qccct tct qtc caaqqga ctt gaat cat qaa aqaqqga ctt at qaact. 1686
|||||

Db 361 TAACCAGTCGGTAGATGCCCTTCTGTCCAGGGACTTGATCATGAAAGAGGACTATGAACT 420

[illegible]

Db 421 TGT TAGTACCAAGCCTACAAGGACCTCAAAAGTCAGACAATTACTAGACACTACTGACAT 480

gy 1747 ccaagggagaaat ttgcgaatttatatcaccaaaat tgaagat acaaaccaat qq; 1806

Db 481 CCAAGGAGAAGAATTTGCCAAAGTTATAGTACAAAAATTGAAAGATAACAAACAAATGGG 540

Oy 1807 telttqaquttlaccedqasalatltlqtuttlitlaqakac atetllaaatlittica 1866
|||||

541 TTTTAACTTAAATTAAGAAATTTTCTTAAATCAATATCTTAAATTACTTCA 600

gy 1867 aaataaaagcaatqtaaqtaqactqt ttttcaaaqaqaaatqtatttcaataa 1916

Db 601 AATAAAAGCATGTAAGTGACGTGTTTTCAAGAAGAAATGTGTTCATAA 650

```

RESULT 4
BM469352
Locus: BM469352 1011 bp mRNA linear F01.05.FEB.2002
DEFINITION: ERBB1 (ErbB1; NTRK1) gene, complete cDNA clone (IMAGE55852)
5', mRNA sequence.
ACCESSION: BM469352
VERSION: BM469352.1 GI:1851894
KEYWORDS: EST.
SOURCE: human.
ORGANISM: Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE:
AUTHOR: Liu M, et al. 1994. J Biol Chem.
TITLE: National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL: unpublished (1999)
COMMENT:
Contact: Robert Strausberg, Ph.D.
Email: rstraub@nih.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LIAM12351 row: e column: 65
High quality sequence stop: 663.

```

```

FEATURES             location/qualifiers
     source            1..1011
                        /organism "Homo sapiens"
                        /db_xref "taxon:9606"
                        /clone="IMAGE:5585260"
                        /date_111 "NTL_MGC_92"
                        /tissue_type "embryonal carcinoma, cell line"
                        /lab_host "DB10B (phage-resistant)"
                        /note "Organ: testis; Vector: pCMV-SPORT6; Site_1: Not1;
                        site_2: Sal1; Cloned unidirectionally; oligo dT primed.
                        Average insert size 2.5 kb. Library enriched for
                        full-length clones and constructed by Life Technologies.
                        Note: this is a NTL MGC Library."

```

BASE COUNT 249 a 275 c 254 g 233 t

matrix: Matrix: 28.5%; Score: 598; DB: 13; Length: 1011;

Db 52 ATCTAATTCAGTGTGTTTTCAGAAATAATGTGTTCATAAAAGGATATTTAT 1

RESULT 5
BE551615/3
LOCUS BE551615 636 bp mRNA linear EST 10-AUG-2000
DEFINITION 7a42q95.x1 NCI-CCAP_G06 Homo sapiens cDNA clone IMAGE:3221434 3' similar to TR-043353-043353 SFRINE/THREONINE KINASE RICK; ; mRNA sequence.

ACCESSION BE551615
VERSION BE551615.1 GI:9793227
KEYWORDS FSI.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 636)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CCAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ccapbs@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNC, send email to: info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 455

FEATURES
source location/qualifiers
1..636
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3221434"
/clone_lib="NCI-CCAP_G06"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pUT3D-P-C (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI-CCAP_G04 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as target in a subtractive hybridization reaction. The driver was PCR amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1257096-1258031, 146906, 1470923, and 1475592-1476742). Subtraction by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 75 a 118 c 135 g 208 t
ORIGIN

Query Match: 26.4%; Score 554; DB 10; Length 636;
Best Local Similarity: 100.0%; Pred. No. 6.6e-119;
Matches: 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Oy 1420 ccttcctgaatacagagtttggatagagcattcttggaatctcaaggagctgcattctg 1479
Db 551 CTGTGCTGAGAAATATACCTGAGAAAGATATATCTGAAATCAAAAGGAGCTGCAATTCG 405
Oy 1480 tgaatcagagagcactccatgctcttgaagataatgaatccactctcaactgcaggaaa 1539
Db 494 TGATCAAGACCACTCCATGCTTCAGCAAATATAAAACACATCACAACGCAGGAAA 435
Oy 1540 ctcaaaactctctagagcttatatagagagcctggatcagaagcaaaaggaagacat 1599
Db 434 CTCAGAAGGTCTGCACCTGCTATAGCCAGCAGGGATCCAGAGCAAAAGGGAACACAT 375
Oy 1600 ttgtgaamaatgacagagagccttcttataccagtcgctgaagatgccttctgtccagga 1659

Db 374 TGTTAAACCAAAATGACAGAAAGCTGCTTCAACCACTGCTGCTAGATTCCTCTGTCCAGGGA 315
Oy 1660 ctctaatcatcaaaadgaatctatgaacttattgaatctcaaaatctcaaaagcctcaaaagt 1719
Db 314 CTGATCAAGAAAGAGGAATATAAATCTGTAGTAAAGGCTTCAAGGACCTCAAAAGT 255
Oy 1720 cagacaaatctatagagcactactgacatccagagagagaaattcttgaagattatattgaa 1779
Db 254 CAGACAATTACTAGACACTACTGACATCCCAAGGAGAAGAAATTGGCAAAAGTTATAGTACA 195
Oy 1780 aaaaattgaagagataacaaacaaatgaattcttcagccttaccgaaatctactgtatctt 1839
Db 194 AAAATGAAAGATACAAACAAATGGAATCTCAGCAATACGCAAAATATCTGTGGTTC 135
Oy 1840 tgaatcagatctcttgaatttacttcagaaataaagaaatctatgaatctatcttctcaga 1899
Db 134 TAGATCACTATCTCTTAAATTTACTTCAAAATAAAAGGATGTAAGTGACIGTCTTCAAGA 75
Oy 1900 agaaatctctcttcaataaaagaaatattatctctctctgtcttgaatttttttatataaa 1959
Db 74 AGAAATGTGTTTCATAAAAGGATAATTATATCTCTGCTGTGGCTTGAATTTTATATAAA 15
Oy 1960 atccatgaatttata 1973
Db 14 ATCCGTGAGATTA 1

RESULT 7
B1257472
LOCUS B1257472 812 bp mRNA linear EST 17-JUL-2001
DEFINITION 602967861P1 NIH-MGC_12 Homo sapiens cDNA clone IMAGE:5107310 5' mRNA sequence.

ACCESSION B1257472
VERSION B1257472.1 GI:14812879
KEYWORDS FSI.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 812)
AUTHORS NIH-MGC <http://www.ncbi.nlm.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: ccapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Inocyte Genomics, Inc.
DNA Sequencing by: Inocyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNC at: info@image.llnl.gov
Plate: L14M11259 rows: 1 column: 15
High quality sequence stop: 806

FEATURES
source location/qualifiers
1..812
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5107310"
/clone_lib="NIH-MGC_12"
/tissue_type="cervical carcinoma cell line"
/lab_host="embryo"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.4 kb. Library prepared by Life Technologies."

BASE COUNT 278 a 170 c 152 g 212 t
ORIGIN

Query Match: 26.3%; Score 551; DB 10; Length 812;
Best Local Similarity: 100.0%; Pred. No. 2.7e-118;

SEQUENCE 8
AA1609-17 1545 bp mRNA linear ESI 16 1057 1996
DEFINITION c1pctcl1 (c) Striatopene hNL neuron (#337243) Homo sapiens cDNA clone.
IMAGE 63044 5'. mRNA sequence.
ACCESSION AA160917
VERSION AA160917.1 GI:1736024
KEYWORDS ENL;
SOURCE Human
ORGANISM Homo sapiens
CHERRY DATA RefSeq; Translators; Translators; Vertebrates; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCES 1. (bases 1 to 645)
AUTHORS Gilchrist, J., Lennon, G., Becker, M., Bonafide, M.F., Chiapelli, R., D'Amico, A., Dietrich, N., Dobson, L., Faveille, X., Gish, W., Hawkins, B., Hoffman, M., Knecht, J., Lady, M., Le, M., Le, N., Mardis, E., Moore, R., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, E., Thierry-Mieg, J., Trevisan, E., Uberwold, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
TITLE Generation and analysis of 200,000 human expressed sequence tags
JOURNAL GENOME RES. 6 (+), 807-828 (1996)
MEDLINE 8744476
COMMENT Contact: William BK Washington University School of Medicine
1444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
tel.: 314.266.1800
fax: 314.266.1810
email: est@watson.wustl.edu
This clone is available royalty free through ENL. For contact the IMAGE Consortium (information.enl.gov) for further information.

Qy	952	TGTGACCAATCCTTTGCAATTAATGTATATGATGTGTGACAGACATGACCTATTATTA	1011
Db	1	TTGTCACCAATCCTTTGCGAGATAAGTATAGTGTGTGCACACAGCAATGCACTGTGTATTA	60
Qy	1012	TGAAGAAAGTTTGGCAATATGATAACCTCCACCGAGCACGATATGATCTCTCTAATAGAAA	1071
Db	61	TGAAGAAAGTTTGGCAATATGATAACCTCCACCGAGCACGATATGATCTCTCTAATAGAAA	120
Qy	1072	TGAATGAGCAACAATCGAATGAAAGACCATCTTCTTGAAGATTTAATGAAGCTTGA	1131
Db	121	TGGATGGGCACAAAAATCCAGATGAAAGACCATCTTCTTGAAGATTTAATGAAGCTTGA	180
Qy	1132	ACCAATTTTGAAGCAATTTGAAGAAATTAACCTTTCTTGAAGCTGTATTCAGCTAAGAA	1191
Db	181	ACCAATTTTGAAGCAATTTGAAGCAATTAACCTTTCTTGAAGCTGTATTCAGCTAAGAA	240
Qy	1192	AACAAAGTTACAGAGTGTTCGAAGTGGCATTCACCTAATGTCACAAAGAAAGAAATGGAA	1251
Db	241	AACAAAGTTACAGAGTGTTCGAAGTGGCATTCACCTAATGTCACAAAGAAAGAAATGGAA	300
Qy	1252	ATCTCTGAACATACCTTGTAAATCATGGTCCACAAAGAGGAATCATGTGGATCTCTCAGCT	1311
Db	301	ATCTCTGAACATACCTTGTAAATCATGGTCCACAAAGAGGAATCATGTGGATCTCTCAGCT	360
Qy	1312	CAATGAAGTAATGATTTCTCTGAAGCTTCAAGTCTCCCTGACAGCTCTCTAAGCAATGA	1371
Db	361	CAATGAAGTAATGATTTCTCTGAAGCTTCAAGTCTCCCTGACAGCTCTCTAAGCAATGA	420
Qy	1372	TTTCTATCTAGAAAAGCTCAAGACTGTTATTTTATGAACTGCAATCACTGCTCTGGAAA	1431
Db	421	TTTCTATCTAGAAAAGCTCAAGACTGTTATTTTATGAACTGCAATCACTGCTCTGGAAA	480
Qy	1432	TCAAGTTGAGTAAGCAATTTCTGAATCTCAAGAGCTGCAATTTCTGAT	1491
Db	481	TCAAGTTGAGTAAGCAATTTCTGAATCTCAAGAGCTGCAATTTCTGAT	540

```

RESULT 1 9
AW960501/c
LOCUS      AW960501                      647 bp      mRNA      linear      EST, Oct 1998
DEFINITION EST 372572 MAGI resequenced, MAGI Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW960501
VERSION    AW960501.1  GI:8150185
KEYWORDS
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo

```

REFERENCE 1 (bases 1 to 647)
AUTHORS Heyde, F., Qi, R., Abernathy, K., Dharap, S., Caspard, R., Gay, C., Holt, J. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnq@tigr.org
Plate: 147
Seq primer: Forward:
Location/Qualifiers
1..647
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGE"
/note="Vector: pBluescriptSKm"
BASE COUNT 165 a 128 c 141 g 23 t
ORIGIN

Query Match 25.4%, Score 532, DB 9, Length 647;
Best Local Similarity 100.0%; Pred. No. 7.7e-114;
Matches 532, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Qy 1396 cgttattttatgaagctgcatcactatctctggaatcacacattgaagatagaccatttc 1455
Db 532 CTGTATTATTTATGAAGCTGCATCAGTCTCTGGAATCACACATTGAAGATAGACCATTTC 473
Qy 1456 tggatctcaaaagggtgcattctgtgatcacaaagccactccatgctcttcagcaataat 1515
Db 472 TGGATCTCAAAGGCTGCATTCTGTGATCACAAAGCCACTCCATGCTCTTCAGCAATAAT 413
Qy 1516 aaatccactctcaactgcaggaaactcagaagcgtctacagccttggtatgcccacgaatg 1575
Db 412 AAATCCACTCTCAACTGCAGGAAACTCAGAAGCGTCTACAGCCTTGGTATGCCACGAATG 353
Qy 1576 gatccaaagcgaagagggaagcattgtgaacaaatgacaaagcctacattacccagatc 1635
Db 352 GATCCAAAGCGAAGGGAAGCATTGTGAACAAATGACAAAGCCTACATTACCCAGATC 293
Qy 1636 gctagatgacctctctgtccagggaacttgatcatgaaagggaactatgaactgtttagtac 1695
Db 292 GCTAGATGACCTCTCTGTCCAGGGAAC TTGATCATGAAAGGGAAC TATGAAC TTGTTAGTAC 233
Qy 1696 caagcctacaaaggaactcaaaaagtcagacaattactagacactactgacatccaaaggaga 1755
Db 232 CAAGCCTACAAAGGAAC TCAAAAAGTCAGACAATTACTAGACACTACTGACATCCAAAGGAGA 173
Qy 1756 ggaatttcgaaggttatgattcgaattttgaaagaaagcaaatggaattcttcagcc 1815
Db 172 AGAATTTCGCAAAAGTTATAGTATCAAAAAATGAAAGATAACAACAAATGGGTCTTCAGCC 113
Qy 1816 ttaacccgaaataacttgattgtgtttctagatcagcctcttttaaaatttacttcaaaaataaaa 1875
Db 112 TTAACCCGAAATAACTTGATTGTGTTTCTAGATCAGCCTCTTTTAAATTTACTTCAAAATAAAAG 53
Qy 1876 catctcaatctcaatgttttttttcaagaaaggaatgtctttatataaaaggaatttta 1927
Db 52 CATCTCAACTCAATGT TTTTTC AAGAAAGGAATGTCTTTATATAAAAGGAATT TTA 1

RESULT 10
BE877822
LOCUS BE877822 870 bp mRNA linear EST 20 OCT 2000
DEFINITION 6U1486392F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3888815,
mRNA sequence.
ACCESSION BE877822
VERSION BE877822.1 GI:19426598

KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 870)
AUTHORS NIH-MGC <http://mgs.cbl.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail@nih.gov
Tissue Procurement: DCTD/DTF/Carida
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Lucyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9669 row: a column: 24
High quality sequence stop: 739.

FEATURES
Source Location/Qualifiers
1..870
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="3888815"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/note="Organ. lung, vector: pCMV SP6B1, Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT 288 a 269 c 170 g 212 t
ORIGIN

Query Match 24.1%, Score 585, DB 13, Length 870;
Best Local Similarity 100.0%; Pred. No. 1e-107;
Matches 595, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

Qy 1153 agagataaacttttcttgaagctgtttatcagctaaagaaagcgttgaagctgttctc 1212
Db 1 AGAGATAAAC TTTCTTGAAGCTGT TATCAGCTAAAGAAAGCAAGCTTACAGAGTGTTC 60
Qy 1213 aagtgcacattacattatgacaaaggaaggaatgaaattatctctctctctctctctct 1272
Db 61 AAGTGCACATTACATTATGACAAAGGAAGGAATGAATTATCTCTCTCTCTCTCTCTCT 120
Qy 1273 tcatggtccacaagagggaatcattggaattctctcagctccatgaaataagtggtctc 1332
Db 121 TCATGCTCCACAAGAGGAATCATCTGGATCTCTCAGCTCCATGAAATAGTGGTCTCTCC 180
Qy 1333 tgaaccttcaaaagctcctctcagctctcgaagacaattgatttttatctcagaagagctc 1392
Db 181 TGAACCTTCAAGGCTCTCTCAGCTCTCGAAGACAATGATTTTATCTCAGAAAGCTCA 240
Qy 1393 aagctgttatttttatgaagctgcacatctgtcctggaattcgaatttgaagatgaat 1452
Db 241 AGACTGTATT TTTATGAAGCTGCATCACTGTCTCTGGAATTCACATTCGCATAGCAAT 300
Qy 1453 tctctggatctcaaaaggtctgatttatatgatacaagaccactctgactcttcaagaa 1512
Db 301 TCTCTGGATCTCAAAGGCTGCATCTGTGATCACAGACATTCATCTCTCTTCAGCAAT 360
Qy 1513 aataaateccactctcaactgcaggaaactcagaacgtctctcagccttggtatagccacga 1572
Db 361 AATAAATCCACTCTCAACTGCAGGAAACTCAGAAGCTCTCCAGCCTGGTATAGCCACGA 420
Qy 1573 gtggatcagaagcaaggaaggaatgtgaaaccaaatgacagaagcctgcttcaagca 1632
Db 421 GTGGATCAGAGCAAGGAAGGACATTGTGAACCAATGACACAAGCCTGCTTAAACA 480
Qy 1633 atcgtatgataacctctctgtccagg 1657

BASE COUNT	166 a	114 c	122 q	204 t	1 other z
ORIGIN					

```

RESULT 13
AI343850/c
LOCUS         AI343850             455 bp      mRNA      linear      EST 02 FEB 1999
DEFINITION    gp07d01.x1 NCI_CGAP_kid5 Homo sapiens cDNA clone IMAGE:191733 3'.
               mRNA sequence.
ACCESSION     AI343850
VERSION       AI343850.1  GI:4081056
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi.
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1. (bases 1 to 455)
AUTHORS      NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
               Tumor Gene Index
JOURNAL       Unpublished (1997)
COMMENT      Contact: Robert Strausberg, Ph.D.
               Email: cgapbs_r@mail.nih.gov
               Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
               Emmert-Buck, M.D., Ph.D.
               cDNA Library Preparation: M. Bento Soares, Ph.D.
               cDNA Library Arrayed by: Greg Lennon, Ph.D.
               DNA Sequencing by: Washington University Genome Sequencing Center
               Clone distribution: NCI-CGAP clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at.

```

```

www.bioinformatics.org/bbip/image/image.html
Insert Length: 930      Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 422.

FEATURES
    source                Location/Qualifiers
                        1..455
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAF:1917313"
                        /clone_lib="NC1-CGAP_kid5"
                        /tissue_type="2 pooled tumors (clear cell type)"
                        /lab_host="DH10B"
                        /note="IMAF: kidney; Vector: pT7T3-Pac (Pharmacia) with
                        a modified polylinker; Site_1, Not I; Site_2, Eco RI; 1st
                        strand cDNA was primed with a Not I + oligo(dT) primer [5'
                        AACCTGGAGAAAGATTCCGCGCGCGCAATATTTTTTTTTTTTT 3'],
                        double-stranded cDNA was ligated to Eco RI adaptors
                        (Pharmacia), digested with Not I, and cloned into the Not I
                        and Eco RI sites of the modified pT7T3 vector. Library
                        went through one round of normalization. Library
                        constructed by Bento Soares and M. Parima Ronaldo."

BASE COUNT      109 a      83 c      95 g      168 t

ORIGIN
    Query Match      21.4%; Score 449; DB 9; Length 455;
    Best Local Similarity 100.0%; Pred. No. 1.3e-94;
    Matches 449; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1481 gatcacaaagacacctccatgctcttcagcaaatatgatatgacttcgaactgcagggaac 1540
Db 455 GATCACAAAGACCACTCCATGCTCTTCAGCAAAATAAIAAAATCCACATCAACTGCCAGGAAC 396
Qy 1541 tcagaacgctctgcagccttggtatagccagcagctggatccagagcdaaaaggaagacatt 1600
Db 395 TCAGAACGCTCTGCAGCCTGCTATAGCCAGCAGTGGATCCAGAGCAAAAGGGAAGACATT 336
Qy 1601 ctgaaccacaatgacacagaagcctgccttaaccagctccttaagtgcctctctcccaagac 1660
Db 335 CTGAACCAAAAGACAGAAAGCCTGCCCTAACACAGTCGCTAGATGCCCTCTCTGTCAGAGAC 275
Qy 1661 ttgatcatgaagaaggaactatgaacttctttagtaccagaagctacagaagacacaaaaatc 1720
Db 275 TTGATCATGAAGAGGACTATGAACCTGTGTAGTACCAAGCCTACAAAGCACTCAAAAGTC 216
Qy 1721 agacataattactagacactactgcacatccaggaaggaaggaatttgaagaagcttatagtacaa 1780
Db 215 AGACAATTACTACAGCACTACTGACATCCAAAGGACAAGAAATTTGCCAAAGTTATAGTACAA 156
Qy 1781 aaattgagaggaatacaacaacaaatgaggtcttcaagccttaccggaataactctatgatttct 1840
Db 155 AAATTGAAAGATAACAACAACAAATGGGCTCTTCAGCGCTACCCGGAAGAAATCTCTGGTTTCT 96
Qy 1841 agatcaccatctctttaaatttacttcaaaaataaagaacatgaagtgactgtttttcaagaa 1900
Db 95 AGATCACCACTCTTAAATTTACTTCAAAAATAAAGCATGTAAGTCACTGTTTTTCAAGAA 36
Qy 1901 gggatggttttgaataaagggtattttata 1924
Db 45 GAAATGTGTTTCATAAAAAGGATAATTATA 7

```

```

RESULT 14
BI820776
LOCUS      B1820776              835 bp      mRNA      linear      EST 04-OCT-2001
DEFINITION 603034.378F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175690 5',
            mRNA sequence.
ACCESSION  BI820776
VERSION    BI820776.1  GI:15932326
KEYWORDS   EST.
SOURCE     human
            Homo sapiens
            ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

DOI: 10.1002/eqm2.12067

Search completed: May 13, 2002, 19:58:11
Job time: 3752 sec

58 qt caqet et qst f rquaaayacacqet qqvtaqvvaat ecqqqqaat qqqqccete 117

[illegible]

```

: RESUL1 2
: US 09/099 041A-1
: Sequence 1, Application US/09099041A
: Patent No. 6 340 576
: GENERAL INFORMATION:
: APPLICANT: Bertin, John
: TITLE OF INVENTION: NOVEL MOLECULES OF THE CARB-RELATED
: TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
: FILE REFERENCE: 07 334/076001
: CURRENT APPLICATION NUMBER: US/09/099 041A
: CURRENT FILING DATE: 1998 06-17
: PRIOR APPLICATION NUMBER: 09/019 942
: PRIOR FILING DATE: 1998 02 06
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID No 1
: LENGTH: 1931
: TYPE: DNA

```

RESULT 3
US-09-069-023-2
; Sequence 2, Application US/09069023A
; Patent No. 6448573
; GENERAL INFORMATION:
; APPLICANT: Nunez, Gabriel

[illegible]

RESULT 5
US-08-700 575-31
: Sequence 31, App. 10/08/08 US/08700575
: Patent No. 5817479

```

RES001 7
US 09/054,298
: Sequence 7. Application US/09054298
: Patent No. 6136954
: GENERAL INFORMATION:
: APPLICANT: Lee, Mu En
: APPLICANT: Haber, Edgar
: APPLICANT: Jain, Mukesh
: APPLICANT: Yet, Shaw Fong
: TITLE OF INVENTION: SMOOTH MUSCLE CELL LIM PROTEIN
: NUMBER OF SEQUENCES: 40
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110 2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows 95
: SOFTWARE: FastSeq for Windows Version 2.0b
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: 09/054,298
: FILING DATE: 02 APR/1998
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: 08/616,368

```



```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

RESULT 12
US 09 468 590-1
: Sequence 1, Application US/09 468 590
: Patent No. 6187563
: GENERAL INFORMATION:
: APPLICANT: Solimena, Michele
: TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
: TITLE OF INVENTION: AUTOANTIGENS OF AUTOIMMUNE DISEASES
: FILE REFERENCE: 101918 200 (OCR 941)
: CURRENT APPLICATION NUMBER: US/09/468,590

```

```

: CURRENT FILING DATE: 1999-08-04
: EARLIER APPLICATION NUMBER: 60/2095,657
: EARLIER FILING DATE: 1998-08-07
: NUMBER OF SEQ. ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ. ID NO 1
:   LENGTH: 7812
:   TYPE: DNA
:   ORGANISM: Human
:   FEATURE:
:     NAME/KEY: CDS
:     LOCATION: (1)...(6879)
:     NAME/KEY: unsure
:     LOCATION: (100)...(102)
:     NAME/KEY: unsure
:     LOCATION: (1021)...(1023)
:     NAME/KEY: unsure
:     LOCATION: (2266)...(2268)
US-09-368-590 1

```

```

Query Match          1.7%; Score 36; 1B 4; Length 7812;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2063 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
      |||
Db 7768 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 7803

```

```

RESULT 14
US-09-592-891A-8
: Sequence 8, Application US/09552891A
: Patent No. 6329174
: GENERAL INFORMATION:
:   APPLICANT: Xiao-Zhou Michelle Wang
:   APPLICANT: Xavier Georges Sarda
:   APPLICANT: Michael David Tomalski
:   APPLICANT: Vincent Paul Mary Wingate
:   TITLE OF INVENTION: Heliothis Glutamate Receptor
:   FILE REFERENCE: A32815-072667-0118
:   CURRENT APPLICATION NUMBER: US/09,592,891A
:   CURRENT FILING DATE: 2001-01-22
:   NUMBER OF SEQ. ID NOS: 14
:   SOFTWARE: FastSeq for Windows Version 3.0
:   SEQ. ID NO 8
:     LENGTH: 1052
:     TYPE: DNA
:     ORGANISM: Heliothis virescens
US-09-592-891A-8

```

```

Query Match          1.7%; Score 35; 1B 4; Length 1052;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2064 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
      |||
Db 1017 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1051

```

```

RESULT 14
US-09-412-102-3
: Sequence 3, Application US/09412102
: Patent No. 6228992
: GENERAL INFORMATION:
:   APPLICANT: JESSEN, HOLLY J
:   APPLICANT: MEYER, TERRY E
:   TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
:   TITLE OF INVENTION: NEMATODES IN PLANTS
:   NUMBER OF SEQUENCES: 10
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)

```

```

: STREET: 3605 Glenwood Ave., Suite 310
: CITY: Raleigh
: STATE: NC
: COUNTRY: US
: ZIP: 27622
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC DOS/MS-DOS
:   SOFTWARE: Patent In Release #1.0, Version #1.0
:   CURRENT APPLICATION DATA:
:     APPLICATION NUMBER: 09/09412102
:     FILING DATE:
:     CLASSIFICATION:
:     PRIOR APPLICATION DATA:
:       APPLICATION NUMBER: 09/217,787
:       FILING DATE:
:     ATTORNEY/AGENT INFORMATION:
:       NAME: Spruill, W. Murray
:       REGISTRATION NUMBER: 32,943
:       REFERENCE/DOCKET NUMBER: 5718-18P1
:   TELECOMMUNICATION INFORMATION:
:     TELEPHONE: 919 420 2202
:     TELEFAX: 919 841 3175
:   INFORMATION PER SEQ. ID NO: 3
:     SEQUENCE CHARACTERISTICS:
:       LENGTH: 1325 base pairs
:       TYPE: nucleic acid
:       STRANDEDNESS: single
:       TOPOLOGY: linear
:     MOLECULE TYPE: cDNA
:     ORIGINAL SOURCE:
:       ORGANISM: Zea mays
:     IMMEDIATE SOURCE:
:       CLONED: P12218
:     FEATURE:
:       NAME/KEY: CDS
:       LOCATION: 126..983
US-09-412-102-3

```

```

Query Match          1.7%; Score 35; 1B 4; Length 1325;
Best Local Similarity 100.0%; Pred. No. 6.5e-06;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2064 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 2098
      |||
Db 1290 tttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1324

```

```

RESULT 14
US-09-217-787-3
: Sequence 3, Application US/09217787
: Patent No. 6284948
: GENERAL INFORMATION:
:   APPLICANT: JESSEN, HOLLY J
:   APPLICANT: MEYER, TERRY E
:   TITLE OF INVENTION: GENES AND METHODS FOR CONTROL OF
:   TITLE OF INVENTION: NEMATODES IN PLANTS
:   NUMBER OF SEQUENCES: 10
:   CORRESPONDENCE ADDRESS:
:     ADDRESSEE: W. Murray Spruill (Alston & Bird, LLP)
:     STREET: 3605 Glenwood Ave., Suite 310
:     CITY: Raleigh
:     STATE: NC
:     COUNTRY: US
:     ZIP: 27622
:   COMPUTER READABLE FORM:
:     MEDIUM TYPE: Floppy disk
:     COMPUTER: IBM PC compatible
:     OPERATING SYSTEM: PC DOS/MS-DOS
:     SOFTWARE: Patent In Release #1.0, Version #1.0
:   CURRENT APPLICATION DATA:

```

Search completed: May 13, 2002, 21:23:39
 Total time: 82.1 sec